```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         June  8, 2003, 15:31:12 ; Search time 7863 Seconds
                                          (without alignments)
                                          11636.681 Million cell updates/sec

Title:          US-09-831-846-1
Perfect score:  3144
Sequence:       1 gcctggctccctctcgctga..........gcggctttggattgcttatg 3144

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                     %
Result               Query
   No.    Score      Match Length DB  ID                       Description
-----------------------------------------------------------------------------
      1     3144     100.0    3144  9  AB033072                 AB033072 Homo sapi
      2     2927      93.1    3320  9  AB056799                 AB056799 Macaca fa
c     3   1424.8      45.3  100426  9  AL591063                 AL591063 Human DNA
      4     1282      40.8  179820  2  AC011679                 AC011679 Homo sapi
      5   1166.6      37.1  159443  2  AC101792                 AC101792 Mus muscu
c     6   1166.6      37.1  303563  2  AC124583                 AC124583 Mus muscu
c     7   1166.6      37.1  308013  2  AC125047                 AC125047 Mus muscu
      8   1151.4      36.6  197893  2  AC127880                 AC127880 Rattus no
c     9    827.4      26.3  197893  2  AC127880                 AC127880 Rattus no
     10    634.6      20.2    1887  6  AX179258                 AX179258 Sequence
     11    634.6      20.2    2549  9  BC003578                 BC003578 Homo sapi
     12      633      20.1    1887  6  AX266997                 AX266997 Sequence
     13      633      20.1    2558  6  AX266995                 AX266995 Sequence
     14      621      19.8    2458  9  BC015581                 BC015581 Homo sapi
     15      608      19.3    1653  6  AX358336                 AX358336 Sequence
     16    587.8      18.7    2612 10   BC023156                 BC023156 Mus muscu
     17    563.8      17.9    3174  9  AB040917                 AB040917 Homo sapi
     18    554.6      17.6    1653  6  AX358338                 AX358338 Sequence
     19    549.2      17.5    3047  9  AK055365                 AK055365 Homo sapi
     20    544.4      17.3    3701  9  AK096627                 AK096627 Homo sapi
c    21    516.4      16.4  180153  2  AC018798                 AC018798 Homo sapi
     22      511      16.3  139072  9  AP000485                 AP000485 Homo sapi
     23      511      16.3  182769  9  AP003176                 AP003176 Homo sapi
     24    477.2      15.2   92797  9  AF038458                 AF038458 Homo sapi
c    25      472      15.0  138538  9  AC011445                 AC011445 Homo sapi
c    26      472      15.0  172090  2  AC018477                 AC018477 Homo sapi
     27      465      14.8   71605  2  AC119332                 AC119332 Rattus no
     28    462.6      14.7    1889  9  BC027475                 BC027475 Homo sapi
c    29    447.8      14.2  173649  2  AC117169                 AC117169 Rattus no
c    30    447.6      14.2  118112  9  HS462C17                 AL033380 Human DNA
     31    446.4      14.2    1683 10   BC004018                 BC004018 Mus muscu
     32      438      13.9  168508  2  AC087141                 AC087141 Mus muscu
     33      438      13.9  198578  2  AC087138                 AC087138 Mus muscu
c    34    432.6      13.8  279589  2  AC073768                 AC073768 Mus muscu
c    35    403.6      12.8  158885  9  CNS01DWT                 AL138498 Human chr
     36    401.4      12.8    4301  9  HSM803436                AL832129 Homo sapi
     37    346.6      11.0    1832  9  BC025310                 BC025310 Homo sapi
c    38    308.6       9.8  138034  2  AC096476                 AC096476 Rattus no
c    39    305.6       9.7  205026  2  AC106165                 AC106165 Rattus no
     40    263.6       8.4    2271  9  BC014678                 BC014678 Homo sapi
c    41    252.2       8.0  100660  2  AC095875                 AC095875 Rattus no
c    42    250.4       8.0   63949  2  AC110498                 AC110498 Mus muscu
     43    213.2       6.8    1307  9  BC014040                 BC014040 Homo sapi
     44      208       6.6    1286  9  BC000207                 BC000207 Homo sapi
     45    204.2       6.5     432  6  AX358333                 AX358333 Sequence

                                 ALIGNMENTS

RESULT 1
AB033072
LOCUS       AB033072                3144 bp    mRNA    linear   PRI 11-NOV-1999
DEFINITION  Homo sapiens mRNA for KIAA1246 protein, partial cds.
ACCESSION   AB033072
VERSION     AB033072.1  GI:6330832
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:hh00149a.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Nagase,T., Ishikawa,K., Kikuno,R., Hirosawa,M., Nomura,N. and
            Ohara,O.
```

```
TITLE       Prediction of the coding sequences of unidentified human genes. XV.
            The complete sequences of 100 new cDNA clones from brain which code
            for large proteins in vitro
  JOURNAL   DNA Res. 6 (5), 337-345 (1999)
  MEDLINE   20039619
REFERENCE   2  (bases 1 to 3144)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1..3144
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="hh00149a"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            1..3144
                     /gene="KIAA1246"
     CDS             <337..2835
                     /gene="KIAA1246"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /product="KIAA1246 protein"
                     /protein_id="BAA86560.1"
                     /db_xref="GI:6330833"
                     /translation="LESVSGGEGCVAEPGSPGAPRSRPRCHPAGGRCCLAQALSDQTM
                     ETLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRL
                     GGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGE
                     DTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNL
                     HQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAP
                     PLSFSFGGNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLI
                     TQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFIT
                     TSQDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSR
                     GGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLI
                     YRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQ
                     SMHSQILGGTMILVIGGIIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNG
                     AQPPPPSSAPAGAPPQGPPKVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAP
                     WRIPPSAPRPKPSLDRLMGAFASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREP
                     LLGPPAARARSLLPLPLEGKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKR
                     SLSVNGMLLPFEESDLVGARGTFGSSEWVMESTV"
BASE COUNT      589 a   1027 c    919 g    609 t
ORIGIN

  Query Match             100.0%;  Score 3144;  DB 9;  Length 3144;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3144;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACACATACACAACCCGGCAGGCTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACACATACACAACCCGGCAGGCTCG 60

Qy         61 TCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGT 120

Qy        121 TCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTT 180

Qy        181 CTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGC 240

Qy        241 CATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCA 300

Qy        301 TCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAA 360

Qy        361 GGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCA 420

Qy        421 GCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTT 480

Qy        481 GGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGT 540

Qy        541 GTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTT 600

Qy        601 GTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATC 660

Qy        661 CACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGG 720

Qy        721 AACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTG 780

Qy        781 CATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAAC 840

Qy        841 CTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAG 900

Qy        901 GACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCG 960

Qy        961 TGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTG 1020

Qy       1021 GATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACC 1080

Qy       1081 TCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCT 1140

Qy       1141 TTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCAC 1200

Qy       1201 TGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGT 1260

Qy       1261 GGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTG 1320

Qy       1321 TGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCG 1380

Qy       1381 GCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCC 1440
```

```
Qy       1441 GATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGAC 1500

Qy       1501 ATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCC 1560

Qy       1561 GGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGC 1620

Qy       1621 ACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACC 1680

Qy       1681 AGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCGGAACGGGCT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCGGAACGGGCT 1740

Qy       1741 GTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCA 1800

Qy       1801 GCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATT 1860

Qy       1861 TACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACT 1920

Qy       1921 GGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACC 1980

Qy       1981 AACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATG 2040

Qy       2041 CACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACG 2100

Qy       2101 CTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCC 2160

Qy       2161 AGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCG 2220

Qy       2221 CCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGC 2280

Qy       2281 AACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGC 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGC 2340

Qy       2341 TCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGACGGGCCCCCTGGAGGATCCCACCCTCC 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGACGGGCCCCCTGGAGGATCCCACCCTCC 2400

Qy       2401 GCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTC 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTC 2460

Qy       2461 AAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGG 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGG 2520

Qy       2521 ACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGG 2580

Qy       2581 GCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGAC 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGAC 2640

Qy       2641 ATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 ATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCT 2700

Qy       2701 CGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCC 2760

Qy       2761 TTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATG 2820

Qy       2821 GAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGGAGAAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGGAGAAG 2880

Qy       2881 GGGAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGG 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGG 2940

Qy       2941 ACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCA 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCA 3000

Qy       3001 CCTCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACT 3060

Qy       3061 GAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAA 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAA 3120

Qy       3121 GCAAGCGGCTTTGGATTGCTTATG 3144
              ||||||||||||||||||||||||
Db       3121 GCAAGCGGCTTTGGATTGCTTATG 3144


RESULT 2
AB056799
LOCUS       AB056799                3320 bp    mRNA    linear   PRI 14-MAR-2001
DEFINITION  Macaca fascicularis brain cDNA clone:QflA-11865, full insert
            sequence.
ACCESSION   AB056799
VERSION     AB056799.1  GI:13365898
KEYWORDS    fis (full insert sequence); oligo capping.
SOURCE      Macaca fascicularis adult male frontal lobe left cDNA to mRNA,
            clone_lib:macaque brain cDNA library QflA clone:QflA-11865.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (sites)
  AUTHORS   Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,
            Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
  TITLE     Isolation of full-length cDNA clones from macaque brain cDNA
            libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3320)
  AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2001) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
```

```
COMMENT     Lab host:     TOP10
            Vector:       pME18S-FL3 (Acc.No. AB009864)
            R.  Site1:    DraIII (CACTGTGTG)
            R.  Site2:    DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR.  The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert.  Libraries
            were constructed by Sugano et al.(University of Tokyo, Institute of
            Medical Science). Custom primer used for sequencing
            ( 5' end primer [CTTCTGCTCTAAAAGCTGCG];
            3' end primer [CGACCTGCAGCTCGAGCACA] ).
FEATURES             Location/Qualifiers
     source          1. .3320
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
                     /clone="QflA-11865"
                     /sex="male"
                     /tissue_type="frontal lobe left"
                     /clone_lib="macaque brain cDNA library QflA"
                     /dev_stage="adult"
                     /note="Host: TOP10; Vector: pME18S-FL3 (Acc.No. AB009864);
                     R. Site1: DraIII (CACTGTGTG); R. Site2: DraIII
                     (CACCATGTG)"
     CDS             571. .2940
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAB39323.1"
                     /db_xref="GI:13365899"
                     /translation="METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLL
                     FVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESL
                     RSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNN
                     LHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIF
                     ARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYF
                     WHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSS
                     RTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATATVEVSIVQLPHLSNSTSRTAPP
                     KSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALAKWSVSKSTPRVK
                     MYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIV
                     GCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRYKVCNHEAPS
                     KMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPPKVVVRNELLDFTASLARASDSSSS
                     SSLGSGEAAGLGRAPWRLPPSAPRPKPSLDRLMGAFASLDLKSQRKEELLDSRTPAGR
                     GAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKRSHSFDMGDFAAAAAGGVVPG
                     GYSPPRRVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSEWVMESTV"
BASE COUNT       628 a   1071 c    967 g    654 t
ORIGIN                 :

  Query Match              93.1%;  Score 2927;  DB 9;  Length 3320;
  Best Local Similarity    96.5%;  Pred. No. 0;
  Matches 3068;  Conservative    0;  Mismatches   75;  Indels   36;  Gaps    6;

Qy          1 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACACATACACAACCCGGCAGGCTCG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         80 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACACATACACAACCC-GCAGGCTCG 138

Qy         61 TCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 TCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGT 198

Qy        121 TCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTAT--TT 178
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||  ||
Db        199 TCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGTGAGGGATTATTATTATTATTATT 258

Qy        179 TTCTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAAC 238
              |||||||||||||||||||||||| |||||||||||||| |||||||||||||||||||
Db        259 TTCTTTAATCTGGAAGAGAAGAGAGCAAGTTGTGCTTTTTCCCCCTTCTTCTTGCTAAAT 318

Qy        239 GCCATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCAC 298
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        319 GCCATGGATATAACTGAATAAGCGGCTCAGGGCTCTCCCCGCGTGGACGTCCGAGGCCAC 378

Qy        299 CATCTGCCTGCATT------------------CGCCGGAGCCGCCGGAGGGTTTAGCTCG 340
              ||||||||||| ||                  ||||||||||||||||||||||||||||
Db        379 CATCTGCCTGCGTTCGCCAGAGCCGCCGGAGCCGCCGGAGCCGCCGGAGGGTTTAGCTCG 438

Qy        341 AGTCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGA 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 AGTCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGA 498

Qy        401 GCCGGCCTCGGTGCCACCCAGCCGGGGGTAGATGCTGCCTCGCCCAGGC-------GCTG 453
              ||||||||||||||| |||||||||||||||||||||||||||||||||        |||
Db        499 GCCGGCCTCGGTGCCTCCCAGCCGGGGGTAGATGCTGCCTCGCCCAGGCACCGTGGCCTG 558

Qy        454 AGTGACCAGACCATGGAGACCCTGCTTGGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCC 513
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        559 AGTGACCAGACCATGGAGACCCTGCTTGGTGGCCTGCTGGCGTTTGGCATGGCGTTTGCC 618

Qy        514 GTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACC 573
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        619 GTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAACCTGTCTGAGTCACTGGGGACC 678

Qy        574 CTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAG 633
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 CTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAG 738

Qy        634.CTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACG 693
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 CTGCGTCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACC 798

Qy        694 GGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTT 753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 GGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTT 858

Qy        754 CTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGG 813
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 CTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGG 918

Qy        814 GAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTG 873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 GAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTG 978

Qy        874 GGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTC 933
              |||||||| ||||||||||||||||||||||||||||||||| ||||||| |||||||||
Db        979 GGCGGCATTGCAGATGAGGCTTTTGAGGACTTCCTGCTGACACTGGAGGACCTGGACCTC 1038

Qy        934 TCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCAC 993
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1039 TCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGGCGCATGGTCAACCTCCAC 1098

Qy        994 CAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTG 1053
              ||||||||||||||||||||||||||||| |||||||| |||||||||||||||||||||
Db       1099 CAGCTGAGCCTGGACCACAACCTGCTGGACCACATCGCTGAGGGCACCTTTGCAGACCTG 1158

Qy       1054 CAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCC 1113
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db       1159 CAGAAACTGGCCCGCCTGGATCTCACCTCCAACCGGCTGCAGAAGCTGCCCCCTGATCCC 1218

Qy       1114 ATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCC 1173
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       1219 ATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCTTTTGCCCCACCCTTGTCC 1278

Qy       1174 TTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTC 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1279 TTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTC 1338

Qy       1234 GAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTC 1293
              ||||| ||||| ||||||||||||||||||||||| |||||||||||||||||||||||||
Db       1339 GAGCGAGACGACGACCTGGAAACCTGTGGCTCCCCCGGGGGCCTCAAGGGTCGCTACTTC 1398

Qy       1294 TGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACAC 1353
              ||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1399 TGGCACGTGCGTGAGGAGGAGTTTGTGTGTGAGCCGCCTCTCATCACCCAGCACACACAC 1458

Qy       1354 AAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCC 1413
```

```
                 ||| |||| ||||||||||||||||||||||||||||||||||||||||| |||||||||
Db          1459 AAGCTGCTAGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATCGGGGACCCC 1518

Qy          1414 AGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACC 1473
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1519 AGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACT 1578

Qy          1474 GCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCC 1533
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1579 GCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCC 1638

Qy          1534 TTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATC 1593
                 ||||||||||||||||||||||| |||||||||||||| |||| ||||||||||||||||
Db          1639 TTCACCTGCATTGCTGCCAATGCCGCCGGAGAGGCCACAGCCACGGTGGAGGTCTCCATC 1698

Qy          1594 GTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTC 1653
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1699 GTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTC 1758

Qy          1654 TCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAG 1713
                 |||||||||||||| ||||||||||||||||||||||| |||||||||||||||||||||
Db          1759 TCAGACATCACTGGTTCCAGCAAGACCAGCCGGGGAGGCGGAGGCAGTGGGGGCGGAGAG 1818

Qy          1714 CCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCC 1773
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db          1819 CCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCAGCT 1878

Qy          1774 CTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTAC 1833
                 |||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db          1879 CTGGCCAAGTGGTCTGTCAGCAAGTCAACACCCCGGGTGAAGATGTACCAGCTGCAGTAC 1938

Qy          1834 AACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTC 1893
                 |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db          1939 AACTGCTCTGACGATGAGGTACTGATCTACAGGATGATCCCAGCCTCCAACAAGGCCTTC 1998

Qy          1894 GTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGAT 1953
                 |||||||||||||||||||| ||||| |||||||||||||||||||||||||||||||||
Db          1999 GTGGTCAACAACCTGGTGTCTGGGACCGGCTACGACTTGTGTGTGCTGGCCATGTGGGAT 2058

Qy          1954 GACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAG 2013
                 ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db          2059 GACACAGCCACAACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAG 2118

Qy          2014 GCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTG 2073
                 |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db          2119 GCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATACTGGGCGGCACCATGATCCTG 2178

Qy          2074 GTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGC 2133
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2179 GTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGC 2238

Qy          2134 TACAAGGTCTGCAACCACGAGGCCCCCAGCAAGATGGCAGCGGCCGTGAGCAATGTGTAC 2193
                 ||||||||||||||||||||||||||||| ||||||||||||||||| ||||||||||||
Db          2239 TACAAGGTCTGCAACCACGAGGCCCCCAGTAAGATGGCAGCGGCCGTCAGCAATGTGTAC 2298

Qy          2194 TCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCG 2253
                 ||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db          2299 TCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCGCCTGCCGGGGCCCCGCCT 2358

Qy          2254 CAGGGCCCGCCGAAGGTGGTGGTGCGCAACGAGCTCCTGGACTTCACCGCCAGCCTGGCC 2313
                 ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db          2359 CAGGGCCCGCCCAAGGTGGTGGTGCGCAACGAGCTCCTGGACTTCACCGCCAGCCTGGCC 2418

Qy          2314 CGCGCCAGTGACTCCTCTTCCTCCAGCTCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGGA 2373
                 ||||| ||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          2419 CGCGCTAGTGACTCCTCTTCCTCCAGCTCCCTGGGCAGTGGGGAGGCCGCGGGGCTGGGA 2478

Qy          2374 CGGGCCCCCTGGAGGATCCCACCCTCCGCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTG 2433
                 ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          2479 CGGGCCCCCTGGAGGCTCCCACCCTCCGCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTG 2538

Qy          2434 ATGGGGGCCTTCGCCTCCCTGGACCTCAAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCC 2493
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2539 ATGGGGGCCTTCGCCTCCCTGGACCTCAAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCC 2598

Qy          2494 AGGACTCCAGCCGGGAGAGGGGCTGGGACGTCGGCCCGGGGCCACCACTCGGACCGAGAG 2553
                 |||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||||
Db          2599 AGGACTCCAGCCGGGAGAGGGGCTGGGACGTCAGCCAGGGGCCACCACTCGGACCGAGAG 2658

Qy          2554 CCACTGCTGGGGCCCCCTGCGGCCCGGGCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGC 2613
                 |||||||||||||| || |||||||||||||||||||||||||||||||| |||||||||
Db          2659 CCACTGCTGGGGCCGCCCGCGGCCCGGGCCAGGAGCCTGCTCCCCTTGCCCTTGGAGGGC 2718

Qy          2614 AAGGCCAAACGCAGCCACTCCTTCGACATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGG 2673
                 |||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||
Db          2719 AAGGCCAAACGCAGTCACTCCTTCGACATGGGGGACTTTGCGGCTGCGGCGGCGGGAGGG 2778

Qy          2674 GTCGTGCCGGGCGGCTACAGTCCTCCTCGGAAGGTCTCGAACATCTGGACGAAGCGCAGC 2733
                 || ||||| |||||||||||||| ||||||| ||||||||||||||||||||||||||||
Db          2779 GTAGTGCCCGGCGGCTACAGTCCCCCTCGGAGGGTCTCGAACATCTGGACGAAGCGCAGC 2838

Qy          2734 CTCTCTGTCAACGGCATGCTCTTGCCCTTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGG 2793
                 ||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |||
Db          2839 CTCTCTGTCAACGGCATGCTCTTGCCCTTCGAGGAGAGTGACCTGGTGGGGGCCCGAGGG 2898

Qy          2794 ACTTTTGGCAGCTCCGAATGGGTGATGGAGAGCACGGTCTAGGTGGGGGTGGGCATGCTC 2853
                 |||||||||||||| || ||||||||||||||||||||||||||||||||||||||||||
Db          2899 ACTTTTGGCAGCTCGGAGTGGGTGATGGAGAGCACGGTCTAGGTGGGGGTGGGCATGCTC 2958

Qy          2854 CCTTTCCTGTGCGCAGGGTGGG---AGAAGGGGAAAGAATCTCACTGGCAAGTGTTTGTG 2910
                 ||||||||| ||||||||||||   |||||||||||||||||||||||||||||||||||
Db          2959 CCTTTCCTGCGCGCAGGGTGGGAGAAGAAGGGGAAAGAATCTCACTGGCAAGTGTTTGTG 3018

Qy          2911 GAGTTTCCATGGTGATGTTTACATCCAGGGACAGTTTCGTCTCCCTGTCAATGGCCTCGT 2970
                 ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db          3019 GAGTTTCCATGGTGATGTTTACATCCAGGGACAGTTTTGTCTCCCTGTCAATGGCCTCGT 3078

Qy          2971 GT-----CCCCCCCTACCCCGCAACACCCACATCACCTCCCCACCACCCGGCCGGGGTGT 3025
                 ||     ||||||| | ||| ||||||||| |||||||||||||||||||||||||||||
Db          3079 GTCCCCCCCCCCCCCAACCCCCAACACCCATATCACCTCCCCACCACCCGGCCGGGGTGT 3138

Qy          3026 GCTCAGGGAATGTGGACTCGCTCAAATGCCGGACTGAGCCCTGAGTGTTTGGAAAGGCGA 3085
                 |||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||
Db          3139 GCTCAGGGAATGCGGACTCGCTCAAATACCGGACTGAGCCCTGAGTGTTTGGAAAGGCGA 3198

Qy          3086 GACTCCGCCTTTCTAATCACAAATGTAGCCTACAAGCAAGCGGCTTTGGATTGCTTATG 3144
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3199 GACTCCGCCTTTCTAATCACAAATGTAGCCTACAAGCAAGCGGCTTTGGATTGCTTATG 3257


RESULT 3
AL591063/c
LOCUS       AL591063              100426 bp    DNA     linear   PRI 19-JUL-2001
DEFINITION  Human DNA sequence from clone RP11-535K1 on chromosome 6, complete
            sequence.
ACCESSION   AL591063
VERSION     AL591063.6  GI:15020551
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100426)
  AUTHORS   Laird,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 25, 2001 this sequence version replaced gi:14787623.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
```

```
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-535K1 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-535K1 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP11-121P10 is at 98427 in this
            sequence. The true right end of clone RP11-552E20 is at 2000 in
            this sequence.
FEATURES             Location/Qualifiers
     source          1. .100426
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-535K1"
                     /clone_lib="RPCI-11.2"
     repeat_region   1. .1026
                     /note="L1M3 repeat: matches 4727. .5817 of consensus"
     repeat_region   1030. .1221
                     /note="L1 repeat: matches 4267. .4460 of consensus"
     repeat_region   1222. .1296
                     /note="Alu repeat: matches 1. .50 of consensus"
     repeat_region   1297. .3135
                     /note="L1 repeat: matches 2379. .4267 of consensus"
     repeat_region   3134. .3536
                     /note="L1ME1 repeat: matches 5186. .5588 of consensus"
     repeat_region   3537. .3675
                     /note="FLAM_A repeat: matches 1. .142 of consensus"
     repeat_region   3677. .4233
                     /note="L1ME1 repeat: matches 5582. .6164 of consensus"
     repeat_region   4523. .5546
                     /note="L1M4 repeat: matches 2429. .3489 of consensus"
     repeat_region   5510. .5956
                     /note="L1MB5 repeat: matches 4780. .5243 of consensus"
     repeat_region   5997. .6289
                     /note="MLT1A1 repeat: matches 102. .365 of consensus"
     repeat_region   6303. .7075
                     /note="L1MB5 repeat: matches 5264. .6022 of consensus"
     repeat_region   7078. .7139
                     /note="MSTA repeat: matches 365. .426 of consensus"
     repeat_region   7147. .7205
                     /note="Alu repeat: matches 244. .302 of consensus"
     repeat_region   7201. .7553
                     /note="MSTA repeat: matches 1. .366 of consensus"
     repeat_region   7555. .7710
                     /note="L1MB6 repeat: matches 6020. .6172 of consensus"
     repeat_region   9233. .9423
                     /note="MIR repeat: matches 22. .219 of consensus"
     repeat_region   9441. .9522
                     /note="41 copies 2 mer tt 63% conserved"
     repeat_region   11776. .11955
                     /note="MIR repeat: matches 59. .235 of consensus"
     repeat_region   15525. .15661
                     /note="L2 repeat: matches 2594. .2744 of consensus"
     repeat_region   15837. .16034
                     /note="MIR repeat: matches 60. .262 of consensus"
     repeat_region   18523. .18696
                     /note="MIR repeat: matches 78. .256 of consensus"
     repeat_region   18748. .19107
                     /note="MER33 repeat: matches 3. .324 of consensus"
     repeat_region   20177. .20290
                     /note="3 copies 38 mer 76% conserved"
     repeat_region   20885. .20965
                     /note="MIR repeat: matches 35. .115 of consensus"
     misc_feature    22278. .22945
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   23265. .23372
                     /note="18 copies 6 mer agagag 64% conserved"
     repeat_region   23279. .23372
                     /note="47 copies 2 mer ag 67% conserved"
     repeat_region   23488. .23613
                     /note="63 copies 2 mer ga 58% conserved"
     repeat_region   23498. .23611
                     /note="19 copies 6 mer gacaga 63% conserved"
     repeat_region   23532. .23603
                     /note="4 copies 18 mer 88% conserved"
     repeat_region   23533. .23604
                     /note="2 copies 36 mer 90% conserved"
     repeat_region   24897. .25274
                     /note="THE1B repeat: matches 1. .364 of consensus"
     repeat_region   25405. .25583
                     /note="MIR repeat: matches 10. .180 of consensus"
     repeat_region   25906. .26082
                     /note="MIR repeat: matches 14. .218 of consensus"
     repeat_region   26339. .26599
                     /note="MIR repeat: matches 2. .262 of consensus"
     repeat_region   26933. .27010
                     /note="13 copies 6 mer tgtgtg 92% conserved"
     repeat_region   27107. .27225
                     /note="MIR repeat: matches 120. .262 of consensus"
     repeat_region   27481. .27801
                     /note="L2 repeat: matches 1904. .2246 of consensus"
     repeat_region   28341. .28546
                     /note="MER20 repeat: matches 9. .218 of consensus"
     repeat_region   28712. .28786
                     /note="LTR33 repeat: matches 10. .87 of consensus"
     repeat_region   29265. .29354
                     /note="LTR33 repeat: matches 110. .198 of consensus"
     repeat_region   29670. .29983
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   30423. .30512
                     /note="L2 repeat: matches 2602. .2695 of consensus"
     repeat_region   30662. .30966
                     /note="AluSx repeat: matches 1. .305 of consensus"
     repeat_region   31122. .31173
                     /note="26 copies 2 mer tg 71% conserved"
     repeat_region   31489. .31573
                     /note="MLT1J repeat: matches 103. .190 of consensus"
     repeat_region   32390. .32444
                     /note="L2 repeat: matches 2651. .2705 of consensus"
     repeat_region   32405. .32618
                     /note="MIR repeat: matches 3. .249 of consensus"
     repeat_region   32729. .32977
                     /note="MIR repeat: matches 14. .262 of consensus"
     repeat_region   33132. .33310
                     /note="L2 repeat: matches 2254. .2449 of consensus"
     repeat_region   33922. .34526
                     /note="L2 repeat: matches 2059. .2672 of consensus"
     repeat_region   34553. .34647
                     /note="L2 repeat: matches 2658. .2749 of consensus"
     repeat_region   34662. .34869
                     /note="MER3 repeat: matches 1. .209 of consensus"
     repeat_region   35504. .35771
                     /note="MIR repeat: matches 2. .257 of consensus"
     repeat_region   35720. .35782
                     /note="L2 repeat: matches 2688. .2750 of consensus"
     repeat_region   36322. .36529
                     /note="MIR repeat: matches 22. .255 of consensus"
     repeat_region   37540. .37724
```

```
                     /note="MIR repeat: matches 4..188 of consensus"
     repeat_region   38100..38197
                     /note="U6 repeat: matches 2..103 of consensus"
     repeat_region   38228..38496
                     /note="L1MB3 repeat: matches 5903..6181 of consensus"
     repeat_region   38815..38998
                     /note="MIR repeat: matches 8..228 of consensus"
     repeat_region   39066..39383
                     /note="MSTD repeat: matches 1..394 of consensus"
     repeat_region   39778..39889
                     /note="MIR repeat: matches 53..165 of consensus"
     repeat_region   40464..40529
                     /note="MIR repeat: matches 51..120 of consensus"
     repeat_region   40657..40862
                     /note="MIR repeat: matches 20..249 of consensus"
     repeat_region   41569..41640
                     /note="36 copies 2 mer gt 66% conserved"
     repeat_region   41737..42201
                     /note="L2 repeat: matches 2203..2701 of consensus"
     repeat_region   42225..42370
                     /note="MIR repeat: matches 116..262 of consensus"
     repeat_region   42733..42853
                     /note="MIR repeat: matches 13..128 of consensus"
     repeat_region   42907..43213
                     /note="L2 repeat: matches 2452..2750 of consensus"
     repeat_region   44325..44471
                     /note="MIR repeat: matches 31..184 of consensus"
     repeat_region   44544..44770
                     /note="MIR repeat: matches 27..262 of consensus"
     repeat_region   45462..45770

  Query Match             45.3%;  Score 1424.8;  DB 9;  Length 100426;
  Best Local Similarity   97.2%;  Pred. No. 3.9e-237;
  Matches 1450;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy         420 AGCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCT 479
               | || | | | |    |  | |  ||||||||||||||||||||||||||||||||||||
Db       63405 ACCCTGAGCTTGTGTTTCTCCCTTCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCT 63346

Qy         480 TGGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTG 539
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63345 TGGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTG 63286

Qy         540 TGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTT 599
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63285 TGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTT 63226

Qy         600 TGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCAT 659
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63225 TGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCAT 63166

Qy         660 CCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAG 719
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63165 CCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAG 63106

Qy         720 GAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCT 779
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63105 GAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCT 63046

Qy         780 GCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAA 839
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63045 GCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAA 62986

Qy         840 CCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGA 899
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62985 CCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGA 62926

Qy         900 GGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCC 959
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62925 GGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCC 62866

Qy         960 GTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCT 1019
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62865 GTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCT 62806

Qy        1020 GGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCAC 1079
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62805 GGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCAC 62746

Qy        1080 CTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGC 1139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62745 CTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGC 62686

Qy        1140 TTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCA 1199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62685 TTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCA 62626

Qy        1200 CTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTG 1259
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62625 CTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTG 62566

Qy        1260 TGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGT 1319
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62565 TGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGT 62506

Qy        1320 GTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGC 1379
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62505 GTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGC 62446

Qy        1380 GGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCC 1439
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62445 GGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCC 62386

Qy        1440 CGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGA 1499
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62385 CGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGA 62326

Qy        1500 CATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGC 1559
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62325 CATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGC 62266

Qy        1560 CGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAG 1619
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62265 CGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAG 62206

Qy        1620 CACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGAC 1679
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62205 CACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGAC 62146

Qy        1680 CAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGC 1739
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62145 CAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGC 62086

Qy        1740 TGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTC 1799
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62085 TGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTC 62026

Qy        1800 AGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGAT 1859
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62025 AGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGAT 61966

Qy        1860 TTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTG 1911
               |||||||         | ||    ||| | |   | ||    |   |||| |
Db       61965 TTACAGGTGAGCAAGCGGCTGTGGCAAAGGCAGGGAGGGTGTCTGGCTGGGG 61914


RESULT 4
AC011679
LOCUS       AC011679               179820 bp    DNA     linear   HTG 26-MAR-2000
DEFINITION  Homo sapiens clone RP11-14G23, WORKING DRAFT SEQUENCE, 16 unordered
            pieces.
ACCESSION   AC011679
VERSION     AC011679.3  GI:7329391
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 179820)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-14G23
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 179820)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 26, 2000 this sequence version replaced gi:6249709.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            ------------- Project Information
                Center project name: L3400
                Center clone name: 14_G_23
            ------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 166032 bases at least Q40
                Consensus quality: 172101 bases at least Q30
                Consensus quality: 174784 bases at least Q20
                Insert size: 182000; agarose-fp
                Insert size: 178320; sum-of-contigs
                Quality coverage: 4.1 in Q20 bases; agarose-fp
                Quality coverage: 4.2 in Q20 bases; sum-of-contigs
            -------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *         1    1960: contig of 1960 bp in length
            *      1961 2060: gap of       100 bp
            *      2061    3217: contig of 1157 bp in length
            *      3218 3317: gap of       100 bp
            *      3318    6117: contig of 2800 bp in length
            *      6118 6217: gap of       100 bp
            *      6218    8883: contig of 2666 bp in length
            *      8884 8983: gap of       100 bp
            *      8984   12533: contig of 3550 bp in length
            *     12534 12633: gap of       100 bp
            *     12634   17388: contig of 4755 bp in length
            *     17389 17488: gap of       100 bp
            *     17489   27094: contig of 9606 bp in length
            *     27095 27194: gap of       100 bp
            *     27195   38372: contig of 11178 bp in length
            *     38373 38472: gap of       100 bp
            *     38473   49446: contig of 10974 bp in length
```

```
            *     49447 49546: gap of       100 bp
            *     49547   60664: contig of 11118 bp in length
            *     60665 60764: gap of       100 bp
            *     60765   74690: contig of 13926 bp in length
            *     74691 74790: gap of       100 bp
            *     74791   89780: contig of 14990 bp in length
            *     89781 89880: gap of       100 bp
            *     89881  109807: contig of 19927 bp in length
            *    109808 109907: gap of       100 bp
            *    109908  128003: contig of 18096 bp in length
            *    128004 128103: gap of       100 bp
            *    128104  148838: contig of 20735 bp in length
            *    148839 148938: gap of       100 bp
            *    148939  179820: contig of 30882 bp in length.
FEATURES             Location/Qualifiers
     source          1..179820
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-14G23"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..1960
                     /note="assembly_fragment"
     misc_feature    2061..3217
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    3318..6117
                     /note="assembly_fragment"
     misc_feature    6218..8883
                     /note="assembly_fragment"
     misc_feature    8984..12533
                     /note="assembly_fragment"
     misc_feature    12634..17388
                     /note="assembly_fragment"
     misc_feature    17489..27094
                     /note="assembly_fragment"
     misc_feature    27195..38372
                     /note="assembly_fragment"
     misc_feature    38473..49446
                     /note="assembly_fragment"
     misc_feature    49547..60664
                     /note="assembly_fragment"
     misc_feature    60765..74690
                     /note="assembly_fragment"
     misc_feature    74791..89780
                     /note="assembly_fragment"
     misc_feature    89881..109807
                     /note="assembly_fragment"
     misc_feature    109908..128003
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    128104..148838
                     /note="assembly_fragment"
     misc_feature    148939..179820
                     /note="assembly_fragment"
BASE COUNT     46444 a  43444 c  42485 g  45943 t   1504 others
ORIGIN

  Query Match             40.8%;  Score 1282;  DB 2;  Length 179820;
  Best Local Similarity   100.0%;  Pred. No. 2e-212;
  Matches 1282;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1863 CAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGG 1922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112240 CAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGG 112299

Qy       1923 CTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACCAA 1982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112300 CTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACCAA 112359

Qy       1983 CATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCA 2042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   112360 CATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCA 112419

Qy     2043 CAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCT 2102
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112420 CAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCT 112479

Qy     2103 GCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG 2162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112480 GCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG 112539

Qy     2163 CAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCC 2222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112540 CAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCC 112599

Qy     2223 TCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGCAA 2282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112600 TCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGCAA 112659

Qy     2283 CGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGCTC 2342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112660 CGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGCTC 112719

Qy     2343 CCTGGGCAGTGGGGAGGCTGCGGGGCTGGACGGGCCCCCTGGAGGATCCCACCCTCCGC 2402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112720 CCTGGGCAGTGGGGAGGCTGCGGGGCTGGACGGGCCCCCTGGAGGATCCCACCCTCCGC 112779

Qy     2403 CCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTCAA 2462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112780 CCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTCAA 112839

Qy     2463 GAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGGAC 2522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112840 GAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGGAC 112899

Qy     2523 GTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGGGC 2582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112900 GTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGGGC 112959

Qy     2583 CAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGACAT 2642
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112960 CAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGACAT 113019

Qy     2643 GGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCTCG 2702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113020 GGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCTCG 113079

Qy     2703 GAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCCTT 2762
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113080 GAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCCTT 113139

Qy     2763 TGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATGGA 2822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113140 TGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATGGA 113199

Qy     2823 GAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGAGAAGGG 2882
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113200 GAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGAGAAGGG 113259

Qy     2883 GAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGAC 2942
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113260 GAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGAC 113319

Qy     2943 AGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCACC 3002
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113320 AGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCACC 113379

Qy     3003 TCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACTGA 3062
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113380 TCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACTGA 113439

Qy     3063 GCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAAGC 3122
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113440 GCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAAGC 113499

Qy     3123 AAGCGGCTTTGGATTGCTTATG 3144
            ||||||||||||||||||||||
Db   113500 AAGCGGCTTTGGATTGCTTATG 113521


RESULT 5
AC101792
LOCUS       AC101792               159443 bp    DNA     linear   HTG 21-AUG-2002
DEFINITION  Mus musculus clone RP24-176G22, WORKING DRAFT SEQUENCE, 20
            unordered pieces.
ACCESSION   AC101792
VERSION     AC101792.2  GI:22381528
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 159443)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP24-176G22
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 159443)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 159443)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Aug 21, 2002 this sequence version replaced gi:17060567.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
```

```
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
    Center: Whitehead Institute/ MIT Center for Genome Research
    Center code: WIBR
    Web site: http://www-seq.wi.mit.edu
    Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
    Center project name: L17407
    Center clone name: 176_G_22
-------------- Summary Statistics
    Sequencing vector: Plasmid; n/a; 100% of reads
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 152592 bases at least Q40
    Consensus quality: 155716 bases at least Q30
    Consensus quality: 156921 bases at least Q20
    Insert size: 157000; agarose-fp
    Insert size: 157543; sum-of-contigs
    Quality coverage: 7.8 in Q20 bases; agarose-fp
    Quality coverage: 7.8 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*          1    1129: contig of 1129 bp in length
*       1130 1229: gap of       100 bp
*       1230    2261: contig of 1032 bp in length
*       2262 2361: gap of       100 bp
*       2362    3703: contig of 1342 bp in length
*       3704 3803: gap of       100 bp
*       3804    5915: contig of 2112 bp in length
*       5916 6015: gap of       100 bp
*       6016    8269: contig of 2254 bp in length
*       8270 8369: gap of       100 bp
*       8370    9416: contig of 1047 bp in length
*       9417 9516: gap of       100 bp
*       9517   12017: contig of 2501 bp in length
*      12018 12117: gap of       100 bp
*      12118   13953: contig of 1836 bp in length
*      13954 14053: gap of       100 bp
*      14054   16946: contig of 2893 bp in length
*      16947 17046: gap of       100 bp
*      17047   20068: contig of 3022 bp in length
*      20069 20168: gap of       100 bp
*      20169   22610: contig of 2442 bp in length
*      22611 22710: gap of       100 bp
*      22711   25193: contig of 2483 bp in length
*      25194 25293: gap of       100 bp
*      25294   28313: contig of 3020 bp in length
*      28314 28413: gap of       100 bp
*      28414   31724: contig of 3311 bp in length
*      31725 31824: gap of       100 bp
*      31825   39932: contig of 8108 bp in length
*      39933 40032: gap of       100 bp
*      40033   50334: contig of 10302 bp in length
*      50335 50434: gap of       100 bp
*      50435   66250: contig of 15816 bp in length
*      66251 66350: gap of       100 bp
*      66351   78871: contig of 12521 bp in length
*      78872 78971: gap of       100 bp
*      78972  118433: contig of 39462 bp in length
*     118434 118533: gap of       100 bp
*     118534  159443: contig of 40910 bp in length.
FEATURES             Location/Qualifiers
     source          1..159443
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="RP24-176G22"
                     /clone_lib="RPCI-24 Male Mouse BAC"
     misc_feature    1..1129
                     /note="assembly_fragment"
     misc_feature    1230..2261
                     /note="assembly_fragment"
     misc_feature    2362..3703
                     /note="assembly_fragment"
     misc_feature    3804..5915
                     /note="assembly_fragment"
     misc_feature    6016..8269
                     /note="assembly_fragment"
     misc_feature    8370..9416
                     /note="assembly_fragment"
     misc_feature    9517..12017
                     /note="assembly_fragment"
     misc_feature    12118..13953
                     /note="assembly_fragment"
     misc_feature    14054..16946
                     /note="assembly_fragment"
     misc_feature    17047..20068
                     /note="assembly_fragment"
     misc_feature    20169..22610
                     /note="assembly_fragment"
     misc_feature    22711..25193
                     /note="assembly_fragment"
     misc_feature    25294..28313
                     /note="assembly_fragment"
     misc_feature    28414..31724
                     /note="assembly_fragment"
     misc_feature    31825..39932
                     /note="assembly_fragment"
     misc_feature    40033..50334
                     /note="assembly_fragment"
     misc_feature    50435..66250
                     /note="assembly_fragment"
     misc_feature    66351..78871
                     /note="assembly_fragment"
     misc_feature    78972..118433
                     /note="assembly_fragment"
     misc_feature    118534..159443
                     /note="assembly_fragment"
BASE COUNT    41724 a  36746 c  35340 g  43731 t   1902 others
ORIGIN

  Query Match             37.1%;  Score 1166.6;  DB 2;  Length 159443;
  Best Local Similarity   86.6%;  Pred. No. 2e-192;
  Matches 1286;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy        426 GGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGG 485
              | ||   |  | ||  |  |||  |||||||||||||||||||||||||| |||||||||||
Db      90878 GTGTCCTTCTTTCCAGGTGCAGAGGCTGAGTGACCAGACCATGGAGACTCTGCTTGGTGG 90937

Qy        486 CCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTG 545
               ||||| || ||||||||||||||||| |||||| |||||||||||||||||| |||||
Db      90938 GCTGCTGGCTTTTGGCATGGCGTTTGCTGTGGTCGATGCCTGCCCCAAGTACTGCGTCTG 90997

Qy        546 CCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACC 605
              ||||||||||||||||||||||||||||||||| |||||||||||||| ||||||||| ||
Db      90998 CCAGAATCTGTCTGAGTCACTGGGGACCCTGTGTCCCTCCAAGGGGCTCCTCTTTGTGCC 91057

Qy        606 CCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACAT 665
              |||||| ||||||||| |||| ||||||||||| ||||| ||||||||||||||||||||
Db      91058 CCCTGACATTGACCGGAGGACGGTGGAGCTGCGTCTGGGTGGCAACTTCATCATCCACAT 91117

Qy        666 CAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACAC 725
                ||||||| ||||||||||||||||| ||||||||||| |||||| ||||||||||||||
Db      91118 TGGCCGCCAAGACTTTGCCAACATGACAGGGCTGGTGGATCTGACCTTGTCCAGGAACAC 91177

Qy        726 CATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCT 785
              ||||||||| ||||||||||| ||||| |||||||| |||||||||||||||||||| ||
Db      91178 CATCAGCCATATCCAGCCCTTCTCCTTCCTGGACCTTGAGAGCCTCCGCTCCCTGCACCT 91237

Qy        786 TGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCA 845
```

```
              ||||||||| ||||| || ||||||||||||||||| |||||||| ||||||||||||||
Db      91238 TGACAGCAACCGGCTACCTAGCCTTGGGGAGGACACACTCCGGGGTCTGGTCAACCTGCA 91297

Qy        846 GCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTT 905
               |||||||| |||||||| ||||||||||| |||||||| ||||| || ||||| |||||
Db      91298 ACACCTTATTGTGAACAATAACCAGCTGGGTGGCATCGCTGATGATGCCTTTGAAGACTT 91357

Qy        906 CCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGA 965
              | ||||||| || || || |||||||| || |||||||||||||| || || || |||||
Db      91358 CTTGCTGACTTTAGAAGACCTGGACCTATCTTACAACAACCTCCACGGACTTCCCTGGGA 91417

Qy        966 CTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCA 1025
               ||||| || ||||||||||||||||| |||||||| ||||||||||| |||||||| ||
Db      91418 TTCCGTACGGCGCATGGTCAACCTCCATCAGCTGAGTCTGGACCACAATCTGCTGGACCA 91477

Qy       1026 CATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAA 1085
              ||| || |||||||| |||||||||||||||||||||||||||||||| ||||| |||||
Db      91478 CATTGCTGAGGGCACTTTTGCAGACCTGCAGAAACTGGCCCGCCTGGACCTCACATCCAA 91537

Qy       1086 TCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGAC 1145
              |||||||||||| || ||||| || ||||||||||||||||||||||||||     | ||
Db      91538 TCGGCTGCAGAAACTTCCCCCGGACCCCATCTTTGCCCGCTCCCAGGCTTCCTTGCTAAC 91597

Qy       1146 AGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAA 1205
               ||||| ||||||||||||||| |||| ||||||||||| || ||||| || ||||||||
Db      91598 TGCCACGCCCTTTGCCCCACCCCTGTCTTTTAGTTTTGGAGGGAACCCGCTGCACTGCAA 91657

Qy       1206 TTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTC 1265
              ||||||||||||||||||||||||||| ||| |||| ||||||||||| |||||||| ||
Db      91658 TTGTGAGCTTCTCTGGCTGCGGAGGCTGGAGAGGGATGATGACCTGGAGACCTGTGGATC 91717

Qy       1266 CCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGA 1325
              ||| ||  | ||||||||||||||||| |||||| | |||||||||||||||||||| ||
Db      91718 CCCCGGAAGTCTCAAGGGTCGCTACTTTTGGCATATTCGTGAGGAGGAGTTTGTGTGTGA 91777

Qy       1326 GCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCAC 1385
              ||||||||||||||||||||||||||||||| |||||||||||||||||||||| |||||
Db      91778 GCCGCCTCTCATCACCCAGCACACACACAAGCTGCTGGTTCTGGAGGGCCAGGCAGCCAC 91837

Qy       1386 ACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGA 1445
               || |||||||| ||||||||||||||||||||  | ||||||||||| |||||||||||
Db      91838 TCTTAAGTGCAAGGCCATTGGGGACCCCAGCCCTTTGATCCACTGGGTTGCCCCCGATGA 91897

Qy       1446 CCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTT 1505
              |||| |||| || ||||| || ||||| ||||| |||||||||||||||||||||||  |
Db      91898 CCGCTTGGTGGGAAACTCTTCCAGGACTGCTGTATATGACAATGGCACCCTGGACATTCT 91957

Qy       1506 CATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGA 1565
               || ||||| ||||||||||||||  |||| |||||||| || |||||||| || |||||
Db      91958 AATTACCACCTCTCAGGACAGTGGACCCTTTACCTGCATCGCAGCCAATGCAGCAGGAGA 92017

Qy       1566 GGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAG 1625
              ||| |||||||  || ||||||||||| || ||||| |||||||| || |||||||| ||
Db      92018 GGCTACGGCCACTGTAGAGGTCTCCATTGTGCAGCTCCCACACCTTAGTAACAGCACTAG 92077

Qy       1626 CCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCG 1685
              ||| |  |||||||||||||| ||||| || |||||||| || |||||||||||||||||
Db      92078 CCGGATGGCACCCCCCAAGTCTCGCCTTTCGGACATCACAGGTTCCAGCAAGACCAGCCG 92137

Qy       1686 GGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCT 1745
              |||||| ||||| |||||||  || ||||||||||||||| |||| ||  ||||||||||
Db      92138 GGGAGGGGGAGGTAGTGGGGCTGGGGAGCCTCCCAAAAGCACCCCAGAGAGGGCTGTGCT 92197

Qy       1746 TGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACC 1805
              |||||| || || |||||||| || |||||||||||||||||||||||||||||||||||
Db      92198 TGTGTCCGATGTTACCACCACATCTGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACC 92257

Qy       1806 CCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAG 1865
              | ||||||||||||| ||||||||||||||||| ||||| |||||||| ||||| |||||
Db      92258 CAGGGTGAAGATGTATCAGCTGCAGTACAACTGTTCTGATGATGAGGTGCTGATCTACAG 92317

Qy       1866 GATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGT 1910
              |     |  || || |     || |     || ||  | |  | |
```

```
Db      92318 GTGAGCCAGAGGCTACCGTGTGGTCAGGTGGGACAGGAGCACGCT 92362

RESULT 6
AC124583/c
LOCUS       AC124583              303563 bp    DNA     linear   HTG 25-JUN-2002
DEFINITION  Mus musculus chromosome UNK clone RP23-123P18, WORKING DRAFT
            SEQUENCE, 14 unordered pieces.
ACCESSION   AC124583
VERSION     AC124583.2  GI:21553290
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 303563)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 303563)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 303563)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Jun 25, 2002 this sequence version replaced gi:21426723.

            ---------------- Genome Center ---------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            -------------- Project Information ------------------
            Center project name: M_BA0123P18
            ----------------------------------------------------
            -------------- Summary Statistics  -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 297108 bases at least Q40
            Consensus quality: 298758 bases at least Q30
            Consensus quality: 299680 bases at least Q20
            Insert size: 211000; agarose-fp
            Insert size: 302263; sum-of-contigs
            Quality coverage: 0.00 in Q20 bases; agarose-fp
            Quality coverage: 10.12 in Q20 bases; sum-of-contigs
            ------------------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1828: contig of 1828 bp in length
            *     1829     1928: gap of unknown length
            *     1929     3716: contig of 1788 bp in length
            *     3717     3816: gap of unknown length
            *     3817     5110: contig of 1294 bp in length
            *     5111     5210: gap of unknown length
            *     5211     9311: contig of 4101 bp in length
            *     9312     9411: gap of unknown length
            *     9412    15660: contig of 6249 bp in length
            *    15661    15760: gap of unknown length
            *    15761    21603: contig of 5843 bp in length
            *    21604    21703: gap of unknown length
```

```
       *      21704      33356: contig of 11653 bp in length
       *      33357      33456: gap of unknown length
       *      33457      44470: contig of 11014 bp in length
       *      44471      44570: gap of unknown length
       *      44571      76854: contig of 32284 bp in length
       *      76855      76954: gap of unknown length
       *      76955     115914: contig of 38960 bp in length
       *     115915     116014: gap of unknown length
       *     116015     144525: contig of 28511 bp in length
       *     144526     144625: gap of unknown length
       *     144626     200283: contig of 55658 bp in length
       *     200284     200383: gap of unknown length
       *     200384     302035: contig of 101652 bp in length
       *     302036     302135: gap of unknown length
       *     302136     303563: contig of 1428 bp in length.
FEATURES             Location/Qualifiers
     source          1..303563
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP23-123P18"
     misc_feature    1..1828
                     /note="assembly_name:Contig24"
     misc_feature    1929..3716
                     /note="assembly_name:Contig25"
     misc_feature    3817..5110
                     /note="assembly_name:Contig26"
     misc_feature    5211..9311
                     /note="assembly_name:Contig27"
     misc_feature    9412..15660
                     /note="assembly_name:Contig28"
     misc_feature    15761..21603
                     /note="assembly_name:Contig29"
     misc_feature    21704..33356
                     /note="assembly_name:Contig30"
     misc_feature    33457..44470
                     /note="assembly_name:Contig31"
     misc_feature    44571..76854
                     /note="assembly_name:Contig32"
     misc_feature    76955..115914
                     /note="assembly_name:Contig33"
     misc_feature    116015..144525
                     /note="assembly_name:Contig34"
     misc_feature    144626..200283
                     /note="assembly_name:Contig35"
     misc_feature    200384..302035
                     /note="assembly_name:Contig36"
     misc_feature    302136..303563
                     /note="assembly_name:Contig9"
BASE COUNT    81812 a  67733 c  68961 g  83749 t   1308 others
ORIGIN

  Query Match             37.1%;  Score 1166.6;  DB 2;  Length 303563;
  Best Local Similarity   86.6%;  Pred. No. 1.9e-192;
  Matches 1286;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy         426 GGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGG 485
               | ||   |  | ||  |  |||  ||||||||||||||||||||||||| |||||||||||
Db      275430 GTGTCCTTCTTTCCAGGTGCAGAGGCTGAGTGACCAGACCATGGAGACTCTGCTTGGTGG 275371

Qy         486 CCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTG 545
                ||||| || ||||||||||||||||| ||||||| |||||||||||||||||| |||||
Db      275370 GCTGCTGGCTTTTGGCATGGCGTTTGCTGTGGTCGATGCCTGCCCCAAGTACTGCGTCTG 275311

Qy         546 CCAGAATCTGTCTGAGTCACTGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACC 605
               |||||||||||||||||||||||||||||||| ||||||||||||||| |||||||| ||
Db      275310 CCAGAATCTGTCTGAGTCACTGGGACCCTGTGTCCCTCCAAGGGGCTCCTCTTTGTGCC 275251

Qy         606 CCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACAT 665
               |||||| ||||||||| |||| |||||||||||| ||||| ||||||||||||||||||||
Db      275250 CCCTGACATTGACCGGAGGACGGTGGAGCTGCGTCTGGGTGGCAACTTCATCATCCACAT 275191

Qy         666 CAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACAC 725
                 ||||||| ||||||||||||||||| ||||||||||| |||||| ||||||||||||||
Db      275190 TGGCCGCCAAGACTTTGCCAACATGACAGGGCTGGTGGATCTGACCTTGTCCAGGAACAC 275131

Qy         726 CATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCT 785
               ||||||||| ||||||||||| ||||| |||||||| |||||||||||||||||||| ||
Db      275130 CATCAGCCATATCCAGCCCTTCTCCTTCCTGGACCTTGAGAGCCTCCGCTCCCTGCACCT 275071

Qy         786 TGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCA 845
               ||||||||| ||||| || |||||||||||||||||| |||||||| ||||||||||||||
Db      275070 TGACAGCAACCGGCTACCTAGCCTTGGGGAGGACACACTCCGGGGTCTGGTCAACCTGCA 275011

Qy         846 GCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTT 905
                |||||||| |||||||| ||||||||||| |||||||| ||||| || ||||| |||||
Db      275010 ACACCTTATTGTGAACAATAACCAGCTGGGTGGCATCGCTGATGATGCCTTTGAAGACTT 274951

Qy         906 CCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGA 965
               | ||||||| || || || |||||||| || |||||||||||||| || || || |||||
Db      274950 CTTGCTGACTTTAGAAGACCTGGACCTATCTTACAACAACCTCCACGGACTTCCCTGGGA 274891

Qy         966 CTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCA 1025
                ||||| || ||||||||||||||||| ||||||| |||||||||||| ||||||| ||
Db      274890 TTCCGTACGGCGCATGGTCAACCTCCATCAGCTGAGTCTGGACCACAATCTGCTGGACCA 274831

Qy        1026 CATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAA 1085
               ||| || |||||||| |||||||||||||||||||||||||||||||| ||||| |||||
Db      274830 CATTGCTGAGGGCACTTTTGCAGACCTGCAGAAACTGGCCCGCCTGGACCTCACATCCAA 274771

Qy        1086 TCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGAC 1145
               |||||||||||| || ||||| || |||||||||||||||||||||||||||     | ||
Db      274770 TCGGCTGCAGAAACTTCCCCCGGACCCCATCTTTGCCCGCTCCCAGGCTTCCTTGCTAAC 274711

Qy        1146 AGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAA 1205
                ||||| |||||||||||||||| |||| |||||||||| || ||||| || ||||||||
Db      274710 TGCCACGCCCTTTGCCCCACCCCTGTCTTTTAGTTTTGGAGGGAACCCGCTGCACTGCAA 274651

Qy        1206 TTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTC 1265
               ||||||||||||||||||||||||||| ||| |||| |||||||||||| ||||||| ||
Db      274650 TTGTGAGCTTCTCTGGCTGCGGAGGCTGGAGAGGGATGATGACCTGGAGACCTGTGGATC 274591

Qy        1266 CCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGA 1325
               ||| ||  | ||||||||||||||||| ||||||| | ||||||||||||||||||| ||
Db      274590 CCCCGGAAGTCTCAAGGGTCGCTACTTTTGGCATATTCGTGAGGAGGAGTTTGTGTGTGA 274531

Qy        1326 GCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCAC 1385
               ||||||||||||||||||||||||||||||| |||||||||||||||||||||| |||||
Db      274530 GCCGCCTCTCATCACCCAGCACACACACAAGCTGCTGGTTCTGGAGGGCCAGGCAGCCAC 274471

Qy        1386 ACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGA 1445
                || |||||||| |||||||||||||||||||||  | ||||||||||| ||||||||||
Db      274470 TCTTAAGTGCAAGGCCATTGGGGACCCCAGCCCTTTGATCCACTGGGTTGCCCCCGATGA 274411

Qy        1446 CCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTT 1505
               |||| |||| || ||||| || ||||| ||||| ||||||||||||||||||||||||  |
Db      274410 CCGCTTGGTGGGAAACTCTTCCAGGACTGCTGTATATGACAATGGCACCCTGGACATTCT 274351

Qy        1506 CATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGA 1565
                || ||||| ||||||||||||||  |||| |||||||| || ||||||||| || |||||
Db      274350 AATTACCACCTCTCAGGACAGTGGACCCTTTACCTGCATCGCAGCCAATGCAGCAGGAGA 274291

Qy        1566 GGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAG 1625
               ||| |||||||  || |||||||||||| || ||||| |||||||| || |||||||| ||
Db      274290 GGCTACGGCCACTGTAGAGGTCTCCATTGTGCAGCTCCCACACCTTAGTAACAGCACTAG 274231

Qy        1626 CCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCG 1685
               ||| |  ||||||||||||||| ||||| || |||||||| || ||||||||||||||||||
Db      274230 CCGGATGGCACCCCCCAAGTCTCGCCTTTCGGACATCACAGGTTCCAGCAAGACCAGCCG 274171

Qy        1686 GGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCT 1745
               |||||| ||||| |||||||  || ||||||||||||||| |||| |   ||||||||||
Db      274170 GGGAGGGGGAGGTAGTGGGGCTGGGGAGCCTCCCAAAAGCACCCCAGAGAGGGCTGTGCT 274111

Qy        1746 TGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACC 1805
               ||||||||| || |||||||| || |||||||||||||||||||||||||||||||||||
```

```
Db     274110 TGTGTCCGATGTTACCACCACATCTGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACC 274051

Qy       1806 CCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAG 1865
              | ||||||||||||| ||||||||||||||||| ||||| |||||||| ||||| |||||
Db     274050 CAGGGTGAAGATGTATCAGCTGCAGTACAACTGTTCTGATGATGAGGTGCTGATCTACAG 273991

Qy       1866 GATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGT 1910
              |     |  || || |     || |     || ||  | |  | |
Db     273990 GTGAGCCAGAGGCTACCGTGTGGTCAGGTGGGACAGGAGCACGCT 273946


RESULT 7
AC125047/c
LOCUS       AC125047              308013 bp    DNA     linear   HTG 29-JUL-2002
DEFINITION  Mus musculus chromosome UNK clone RP23-354G14, WORKING DRAFT
            SEQUENCE, 7 unordered pieces.
ACCESSION   AC125047
VERSION     AC125047.2  GI:22002214
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 308013)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 308013)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 308013)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUL-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Jul 29, 2002 this sequence version replaced gi:21490485.

            ---------------- Genome Center ---------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            -------------- Project Information -------------------
            Center project name: M_BA0354G14
            ------------------------------------------------------
            -------------- Summary Statistics  -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 303857 bases at least Q40
            Consensus quality: 305426 bases at least Q30
            Consensus quality: 306387 bases at least Q20
            Insert size: 196000; agarose-fp
            Insert size: 307413; sum-of-contigs
            Quality coverage: 0.00 in Q20 bases; agarose-fp
            Quality coverage: 9.67 in Q20 bases; sum-of-contigs
            -----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1     12436: contig of 12436 bp in length
            *      12437     12536: gap of unknown length
            *      12537     40835: contig of 28299 bp in length
            *      40836     40935: gap of unknown length
            *      40936     58095: contig of 17160 bp in length
            *      58096     58195: gap of unknown length
            *      58196    142400: contig of 84205 bp in length
            *     142401    142500: gap of unknown length
            *     142501    202912: contig of 60412 bp in length
            *     202913    203012: gap of unknown length
            *     203013    306579: contig of 103567 bp in length
            *     306580    306679: gap of unknown length
            *     306680    308013: contig of 1334 bp in length.
FEATURES             Location/Qualifiers
     source          1..308013
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP23-354G14"
     misc_feature    1..12436
                     /note="assembly_name:Contig10"
     misc_feature    12537..40835
                     /note="assembly_name:Contig11"
     misc_feature    40936..58095
                     /note="assembly_name:Contig12"
     misc_feature    58196..142400
                     /note="assembly_name:Contig13"
     misc_feature    142501..202912
                     /note="assembly_name:Contig14"
     misc_feature    203013..306579
                     /note="assembly_name:Contig15"
     misc_feature    306680..308013
                     /note="assembly_name:Contig8"
BASE COUNT    90720 a  61748 c  61427 g  93518 t    600 others
ORIGIN

  Query Match             37.1%;  Score 1166.6;  DB 2;  Length 308013;
  Best Local Similarity   86.6%;  Pred. No. 1.9e-192;
  Matches 1286;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy        426 GGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGG 485
              | ||   |  | ||  |  |||  ||||||||||||||||||||||||| |||||||||||
Db      26361 GTGTCCTTCTTTCCAGGTGCAGAGGCTGAGTGACCAGACCATGGAGACTCTGCTTGGTGG 26302

Qy        486 CCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTG 545
               ||||| || ||||||||||||||||| |||||||| ||||||||||||||||| |||||
Db      26301 GCTGCTGGCTTTGGCATGGCGTTTGCTGTGGTCGATGCCTGCCCCAAGTACTGCGTCTG 26242

Qy        546 CCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACC 605
              ||||||||||||||||||||||||||||||||| |||||||||||||| |||||||| ||
Db      26241 CCAGAATCTGTCTGAGTCACTGGGGACCCTGTGTCCCTCCAAGGGGCTCCTCTTTGTGCC 26182

Qy        606 CCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACAT 665
              |||||| ||||||||| |||| ||||||||||| ||||| ||||||||||||||||||||
Db      26181 CCCTGACATTGACCGGAGGACGGTGGAGCTGCGTCTGGGTGGCAACTTCATCATCCACAT 26122

Qy        666 CAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACAC 725
                ||||||| ||||||||||||||||| ||||||||||| |||||| |||||||||||||
Db      26121 TGGCCGCCAAGACTTTGCCAACATGACAGGGCTGGTGGATCTGACCTTGTCCAGGAACAC 26062

Qy        726 CATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCT 785
              ||||||||| ||||||||||| ||||| |||||||| |||||||||||||||||||| ||
Db      26061 CATCAGCCATATCCAGCCCTTCTCCTTCCTGGACCTTGAGAGCCTCCGCTCCCTGCACCT 26002

Qy        786 TGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCA 845
              ||||||||| ||||| || ||||||||||||||||| |||||||| ||||||||||||||
Db      26001 TGACAGCAACCGGCTACCTAGCCTTGGGGAGGACACACTCCGGGGTCTGGTCAACCTGCA 25942

Qy        846 GCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTT 905
               |||||||| |||||||| ||||||||||| |||||||| ||||| || ||||| |||||
Db      25941 ACACCTTATTGTGAACAATAACCAGCTGGGTGGCATCGCTGATGATGCCTTTGAAGACTT 25882

Qy        906 CCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGA 965
              | ||||||| || || || |||||||| || |||||||||||||| || || || |||||
Db      25881 CTTGCTGACTTTAGAAGACCTGGACCTATCTTACAACAACCTCCACGGACTTCCCTGGGA 25822

Qy        966 CTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCA 1025
```

```
               ||||| || ||||||||||||||||| |||||||| ||||||||||| |||||||| ||
Db      25821 TTCCGTACGGCGCATGGTCAACCTCCATCAGCTGAGTCTGGACCACAATCTGCTGGACCA 25762

Qy       1026 CATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAA 1085
              ||| || |||||||| |||||||||||||||||||||||||||||||| ||||| |||||
Db      25761 CATTGCTGAGGGCACTTTTGCAGACCTGCAGAAACTGGCCCGCCTGGACCTCACATCCAA 25702

Qy       1086 TCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGAC 1145
              |||||||||||| || ||||| || ||||||||||||||||||||||||||     | ||
Db      25701 TCGGCTGCAGAAACTTCCCCCGGACCCCATCTTTGCCCGCTCCCAGGCTTCCTTGCTAAC 25642

Qy       1146 AGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGTAACCCACTTCACTGCAA 1205
               ||||| ||||||||||||||| |||| ||||||||||| || ||||| || ||||||||
Db      25641 TGCCACGCCCTTTGCCCCACCCCTGTCTTTTAGTTTTGGAGGGAACCCGCTGCACTGCAA 25582

Qy       1206 TTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTC 1265
              ||||||||||||||||||||||||||| ||| |||| ||||||||||| |||||||| ||
Db      25581 TTGTGAGCTTCTCTGGCTGCGGAGGCTGGAGAGGGATGATGACCTGGAGACCTGTGGATC 25522

Qy       1266 CCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGA 1325
              ||| ||  | ||||||||||||||||| |||||| | |||||||||||||||||||| ||
Db      25521 CCCCGGAAGTCTCAAGGGTCGCTACTTTTGGCATATTCGTGAGGAGGAGTTTGTGTGTGA 25462

Qy       1326 GCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCAC 1385
              ||||||||||||||||||||||||||||||| |||||||||||||||||||||| |||||
Db      25461 GCCGCCTCTCATCACCCAGCACACACACAAGCTGCTGGTTCTGGAGGGCCAGGCAGCCAC 25402

Qy       1386 ACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGA 1445
               || |||||||| ||||||||||||||||||||  | ||||||||||| |||||||||||
Db      25401 TCTTAAGTGCAAGGCCATTGGGGACCCCAGCCCTTTGATCCACTGGGTTGCCCCCGATGA 25342

Qy       1446 CCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTT 1505
              |||| |||| || ||||| || ||||| ||||| |||||||||||||||||||||||  |
Db      25341 CCGCTTGGTGGGAAACTCTTCCAGGACTGCTGTATATGACAATGGCACCCTGGACATTCT 25282

Qy       1506 CATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGA 1565
               || ||||| ||||||||||||||  |||| |||||||| || |||||||| || |||||
Db      25281 AATTACCACCTCTCAGGACAGTGGACCCTTTACCTGCATCGCAGCCAATGCAGCAGGAGA 25222

Qy       1566 GGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAG 1625
              ||| |||||||  || ||||||||||| || ||||| |||||||| || |||||||| ||
Db      25221 GGCTACGGCCACTGTAGAGGTCTCCATTGTGCAGCTCCCACACCTTAGTAACAGCACTAG 25162

Qy       1626 CCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCG 1685
              ||| |  |||||||||||||| ||||| || |||||||| || |||||||||||||||||
Db      25161 CCGGATGGCACCCCCCAAGTCTCGCCTTTCGGACATCACAGGTTCCAGCAAGACCAGCCG 25102

Qy       1686 GGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCT 1745
              |||||| ||||| |||||||  || ||||||||||||||| |||| ||  ||||||||||
Db      25101 GGGAGGGGGAGGTAGTGGGGCTGGGGAGCCTCCCAAAAGCACCCCAGAGAGGGCTGTGCT 25042

Qy       1746 TGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACC 1805
              |||||| || || |||||||| || |||||||||||||||||||||||||||||||||||
Db      25041 TGTGTCCGATGTTACCACCACATCTGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACC 24982

Qy       1806 CCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAG 1865
              | ||||||||||||| ||||||||||||||||| ||||| |||||||| ||||| |||||
Db      24981 CAGGGTGAAGATGTATCAGCTGCAGTACAACTGTTCTGATGATGAGGTGCTGATCTACAG 24922

Qy       1866 GATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGT 1910
              |     |  || || |     || |     || ||  | |  | |
Db      24921 GTGAGCCAGAGGCTACCGTGTGGTCAGGTGGGACAGGAGCACGCT 24877

RESULT 8
AC127880
LOCUS       AC127880                197893 bp    DNA     linear   HTG 19-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-5008, *** SEQUENCING IN PROGRESS ***,
            63 unordered pieces.
ACCESSION   AC127880
VERSION     AC127880.1  GI:21908317
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 197893)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 197893)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GXPO
                Center clone name: CH230-5008
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 143066 bases at least Q40
                Consensus quality: 151342 bases at least Q30
                Consensus quality: 157286 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 63 contigs. The true order of the pieces
```

```
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1      1082: contig of 1082 bp in length
*     1083      1182: gap of unknown length
*     1183      2644: contig of 1462 bp in length
*     2645      2744: gap of unknown length
*     2745      3881: contig of 1137 bp in length
*     3882      3981: gap of unknown length
*     3982      5343: contig of 1362 bp in length
*     5344      5443: gap of unknown length
*     5444      6742: contig of 1299 bp in length
*     6743      6842: gap of unknown length
*     6843      8343: contig of 1501 bp in length
*     8344      8443: gap of unknown length
*     8444     10196: contig of 1753 bp in length
*    10197     10296: gap of unknown length
*    10297     11522: contig of 1226 bp in length
*    11523     11622: gap of unknown length
*    11623     13415: contig of 1793 bp in length
*    13416     13515: gap of unknown length
*    13516     15170: contig of 1655 bp in length
*    15171     15270: gap of unknown length
*    15271     17014: contig of 1744 bp in length
*    17015     17114: gap of unknown length
*    17115     18688: contig of 1574 bp in length
*    18689     18788: gap of unknown length
*    18789     20090: contig of 1302 bp in length
*    20091     20190: gap of unknown length
*    20191     22262: contig of 2072 bp in length
*    22263     22362: gap of unknown length
*    22363     24086: contig of 1724 bp in length
*    24087     24186: gap of unknown length
*    24187     25869: contig of 1683 bp in length
*    25870     25969: gap of unknown length
*    25970     28059: contig of 2090 bp in length
*    28060     28159: gap of unknown length
*    28160     30192: contig of 2033 bp in length
*    30193     30292: gap of unknown length
*    30293     32224: contig of 1932 bp in length
*    32225     32324: gap of unknown length
*    32325     34517: contig of 2193 bp in length
*    34518     34617: gap of unknown length
*    34618     35832: contig of 1215 bp in length
*    35833     35932: gap of unknown length
*    35933     38365: contig of 2433 bp in length
*    38366     38465: gap of unknown length
*    38466     39651: contig of 1186 bp in length
*    39652     39751: gap of unknown length
*    39752     42696: contig of 2945 bp in length
*    42697     42796: gap of unknown length
*    42797     44523: contig of 1727 bp in length
*    44524     44623: gap of unknown length
*    44624     47132: contig of 2509 bp in length
*    47133     47232: gap of unknown length
*    47233     49315: contig of 2083 bp in length
*    49316     49415: gap of unknown length
*    49416     52379: contig of 2964 bp in length
*    52380     52479: gap of unknown length
*    52480     54484: contig of 2005 bp in length
*    54485     54584: gap of unknown length
*    54585     57588: contig of 3004 bp in length
*    57589     57688: gap of unknown length
*    57689     59983: contig of 2295 bp in length
*    59984     60083: gap of unknown length
*    60084     62833: contig of 2750 bp in length
*    62834     62933: gap of unknown length
*    62934     65460: contig of 2527 bp in length
*    65461     65560: gap of unknown length
*    65561     67597: contig of 2037 bp in length
*    67598     67697: gap of unknown length
*    67698     71010: contig of 3313 bp in length
*    71011     71110: gap of unknown length
*    71111     73099: contig of 1989 bp in length
*    73100     73199: gap of unknown length
*    73200     76659: contig of 3460 bp in length
*    76660     76759: gap of unknown length
*    76760     79328: contig of 2569 bp in length
*    79329     79428: gap of unknown length
*    79429     82950: contig of 3522 bp in length
*    82951     83050: gap of unknown length
*    83051     85337: contig of 2287 bp in length
*    85338     85437: gap of unknown length
*    85438     88632: contig of 3195 bp in length
*    88633     88732: gap of unknown length
*    88733     92607: contig of 3875 bp in length
*    92608     92707: gap of unknown length
*    92708     95823: contig of 3116 bp in length
*    95824     95923: gap of unknown length
*    95924     98417: contig of 2494 bp in length
*    98418     98517: gap of unknown length
*    98518    102227: contig of 3710 bp in length
*   102228    102327: gap of unknown length
*   102328    106214: contig of 3887 bp in length
*   106215    106314: gap of unknown length
*   106315    109534: contig of 3220 bp in length
*   109535    109634: gap of unknown length
*   109635    113077: contig of 3443 bp in length
*   113078    113177: gap of unknown length
*   113178    116869: contig of 3692 bp in length
*   116870    116969: gap of unknown length
*   116970    121793: contig of 4824 bp in length
*   121794    121893: gap of unknown length
*   121894    126597: contig of 4704 bp in length
*   126598    126697: gap of unknown length
*   126698    131888: contig of 5191 bp in length
*   131889    131988: gap of unknown length
*   131989    136453: contig of 4465 bp in length
*   136454    136553: gap of unknown length
*   136554    141386: contig of 4833 bp in length
*   141387    141486: gap of unknown length
*   141487    146790: contig of 5304 bp in length
*   146791    146890: gap of unknown length
*   146891    151544: contig of 4654 bp in length
*   151545    151644: gap of unknown length
*   151645    157304: contig of 5660 bp in length
*   157305    157404: gap of unknown length

  Query Match             36.6%;  Score 1151.4;  DB 2;  Length 197893;
  Best Local Similarity   86.6%;  Pred. No. 8.5e-190;
  Matches 1269;  Conservative    0;  Mismatches  196;  Indels    0;  Gaps    0;

Qy        426 GGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGG 485
              | ||   | || ||  |  |||  || ||||||||||||||||||||| |||||||| ||
Db      42887 GTGTCCTTCCTTCCAGGTGCAGAGGCCGAGTGACCAGACCATGGAGACTCTGCTTGGCGG 42946

Qy        486 CCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTG 545
               ||||| || ||||||||||||||||| |||||||| ||||||||||||||||| |||||
Db      42947 GCTGCTGGCTTTTGGCATGGCGTTTGCTGTGGTCGATGCCTGCCCCAAGTACTGCGTCTG 43006

Qy        546 CCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACC 605
              ||||||||||||||||||||||||||||||||| || ||||||||||| ||||| || ||
Db      43007 CCAGAATCTGTCTGAGTCACTGGGGACCCTGTGTCCTTCCAAGGGGCTCCTCTTCGTGCC 43066

Qy        606 CCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACAT 665
              |||||| ||||||||| |||||||||||||||| ||||| ||||||||||||||||||||
Db      43067 CCCTGACATTGACCGGAGGACAGTGGAGCTGCGTCTGGGTGGCAACTTCATCATCCACAT 43126

Qy        666 CAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACAC 725
                ||||||| ||||||||||||||||| ||||| ||||| |||||| |||| ||||||||
Db      43127 TGGCCGCCAAGACTTTGCCAACATGACAGGGCTAGTGGATCTGACCTTGTCTAGGAACAC 43186

Qy        726 CATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCT 785
```

```
              ||||||||||||||||||||| ||||| |||||||| |||||||||| ||||||||| ||
Db      43187 CATCAGCCACATCCAGCCCTTCTCCTTCCTGGACCTTGAGAGCCTCCGTTCCCTGCACCT 43246

Qy        786 TGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCA 845
              ||||||||| ||||| || ||||||||||||||||| ||||||||||| |||||||||||
Db      43247 TGACAGCAACCGGCTACCTAGCCTTGGGGAGGACACACTCCGGGGCCTAGTCAACCTGCA 43306

Qy        846 GCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTT 905
              ||||||||| |||||||| || |||||||| ||||| || ||||| || ||||| |||||
Db      43307 GCACCTTATTGTGAACAATAATCAGCTGGGTGGCATTGCTGATGATGCCTTTGAAGACTT 43366

Qy        906 CCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGA 965
              |||| |||| || || ||| ||||||| || |||||||||||||| || || || |||||
Db      43367 CCTGTTGACTTTAGAAGATTTGGACCTATCTTACAACAACCTCCACGGTCTTCCCTGGGA 43426

Qy        966 CTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCA 1025
               || ||  | ||||||||||||||||| |||||||| || ||||||||| ||||||| ||
Db      43427 TTCTGTAAGGCGCATGGTCAACCTCCATCAGCTGAGTCTCGACCACAACTTGCTGGACCA 43486

Qy       1026 CATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAA 1085
              |||||| ||||| ||||| ||||||||||||||||||||||||||||| ||||| |||||
Db      43487 CATCGCTGAGGGTACCTTCGCAGACCTGCAGAAACTGGCCCGCCTGGACCTCACGTCCAA 43546

Qy       1086 TCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGAC 1145
              ||||||||||||||| ||||| || ||||||||||||||||||||||| ||     | ||
Db      43547 TCGGCTGCAGAAGCTTCCCCCAGACCCCATCTTTGCCCGCTCCCAGGCCTCCTTGCTCAC 43606

Qy       1146 AGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAA 1205
               ||||||||||||||||||||| |||||||||||||||| || |||||||| ||||||||
Db      43607 TGCCACACCCTTTGCCCCACCCCTGTCCTTTAGTTTTGGAGGGAACCCACTGCACTGCAA 43666

Qy       1206 TTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTC 1265
              ||||||||||||||||||||| ||||| ||| |||| || |||||||| |||||||| ||
Db      43667 TTGTGAGCTTCTCTGGCTGCGAAGGCTGGAGAGGGATGACGACCTGGAGACCTGTGGATC 43726

Qy       1266 CCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGA 1325
              ||| || || |||||||||||||| || |||||| | || ||||||||||||||||| ||
Db      43727 CCCTGGAGGTCTCAAGGGTCGCTATTTTTGGCATATTCGCGAGGAGGAGTTTGTGTGTGA 43786

Qy       1326 GCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCAC 1385
              |||||| |||||||||||||||||||||||| |||||||||||||||||||||| |||||
Db      43787 GCCGCCCCTCATCACCCAGCACACACACAAGCTGCTGGTTCTGGAGGGCCAGGCAGCCAC 43846

Qy       1386 ACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGA 1445
               || |||||||| ||||||||||||||||||||  | ||||||||||| |||||||||||
Db      43847 TCTTAAGTGCAAGGCCATTGGGGACCCCAGCCCTTTGATCCACTGGGTTGCCCCCGATGA 43906

Qy       1446 CCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTT 1505
              |||| ||||||| ||||| || ||||| || || |||||||| ||||| ||||||||  |
Db      43907 CCGCTTGGTAGGAAACTCTTCCAGGACTGCCGTATATGACAACGGCACTCTGGACATTCT 43966

Qy       1506 CATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGA 1565
               || ||||| ||||||||||| ||  |||||||||||||||| ||||| || || |||||
Db      43967 AATTACCACCTCTCAGGACAGCGGCCCCTTCACCTGCATTGCAGCCAACGCGGCAGGAGA 44026

Qy       1566 GGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAG 1625
              ||| || ||||  || |||||||| ||||| ||||| ||||||||||| |||||||| ||
Db      44027 GGCTACCGCCACCGTAGAGGTCTCTATCGTGCAGCTCCCACACCTCAGTAACAGCACTAG 44086

Qy       1626 CCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCG 1685
              ||| |  || || |||||||| ||||||||||||||||| || |||||||||||||| ||
Db      44087 CCGGATGGCGCCGCCCAAGTCTCGCCTCTCAGACATCACGGGTTCCAGCAAGACCAGTCG 44146

Qy       1686 GGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCT 1745
              |||||| ||||| || ||||  || ||||||||||||||| |||| ||  ||||||||||
Db      44147 GGGAGGGGGAGGTAGCGGGGCTGGGGAGCCTCCCAAAAGCACCCCAGAGAGGGCTGTGCT 44206

Qy       1746 TGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACC 1805
              ||||||||| ||||||||||| || ||||||||||||||||||||||||||||| |||||
Db      44207 TGTGTCTGATGTGACCACCACATCTGCCCTGGTCAAGTGGTCTGTCAGCAAGTCGGCACC 44266

Qy       1806 CCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAG 1865
              | ||||||||||||| |||||||||||||||||||| ||||| |||||||||||||||||
Db      44267 CAGGGTGAAGATGTATCAGCTGCAGTACAACTGCTCCGACGACGAGGTACTGATTTACAG 44326

Qy       1866 GATGATCCCAGCCTCCAACAAGGCC 1890
              |     |  || || ||    || |
Db      44327 GTGAGCCAGAGGCTGCAGTGTGGTC 44351


RESULT 9
AC127880/c
LOCUS       AC127880                197893 bp    DNA     linear   HTG 19-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-5O08, *** SEQUENCING IN PROGRESS ***,
            63 unordered pieces.
ACCESSION   AC127880
VERSION     AC127880.1  GI:21908317
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 197893)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 197893)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
```

```
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: GXPO
    Center clone name: CH230-5008
-------------- Summary Statistics
    Sequencing vector: Plasmid;
    Chemistry: Dye-terminator Big Dye: 100% of reads
    Assembly program: Phrap; version 0.990329
    Consensus quality: 143066 bases at least Q40
    Consensus quality: 151342 bases at least Q30
    Consensus quality: 157286 bases at least Q20
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1082: contig of 1082 bp in length
*     1083     1182: gap of unknown length
*     1183     2644: contig of 1462 bp in length
*     2645     2744: gap of unknown length
*     2745     3881: contig of 1137 bp in length
*     3882     3981: gap of unknown length
*     3982     5343: contig of 1362 bp in length
*     5344     5443: gap of unknown length
*     5444     6742: contig of 1299 bp in length
*     6743     6842: gap of unknown length
*     6843     8343: contig of 1501 bp in length
*     8344     8443: gap of unknown length
*     8444    10196: contig of 1753 bp in length
*    10197    10296: gap of unknown length
*    10297    11522: contig of 1226 bp in length
*    11523    11622: gap of unknown length
*    11623    13415: contig of 1793 bp in length
*    13416    13515: gap of unknown length
*    13516    15170: contig of 1655 bp in length
*    15171    15270: gap of unknown length
*    15271    17014: contig of 1744 bp in length
*    17015    17114: gap of unknown length
*    17115    18688: contig of 1574 bp in length
*    18689    18788: gap of unknown length
*    18789    20090: contig of 1302 bp in length
*    20091    20190: gap of unknown length
*    20191    22262: contig of 2072 bp in length
*    22263    22362: gap of unknown length
*    22363    24086: contig of 1724 bp in length
*    24087    24186: gap of unknown length
*    24187    25869: contig of 1683 bp in length
*    25870    25969: gap of unknown length
*    25970    28059: contig of 2090 bp in length
*    28060    28159: gap of unknown length
*    28160    30192: contig of 2033 bp in length
*    30193    30292: gap of unknown length
*    30293    32224: contig of 1932 bp in length
*    32225    32324: gap of unknown length
*    32325    34517: contig of 2193 bp in length
*    34518    34617: gap of unknown length
*    34618    35832: contig of 1215 bp in length
*    35833    35932: gap of unknown length
*    35933    38365: contig of 2433 bp in length
*    38366    38465: gap of unknown length
*    38466    39651: contig of 1186 bp in length
*    39652    39751: gap of unknown length
*    39752    42696: contig of 2945 bp in length
*    42697    42796: gap of unknown length
*    42797    44523: contig of 1727 bp in length
*    44524    44623: gap of unknown length
*    44624    47132: contig of 2509 bp in length
*    47133    47232: gap of unknown length
*    47233    49315: contig of 2083 bp in length
*    49316    49415: gap of unknown length
*    49416    52379: contig of 2964 bp in length
*    52380    52479: gap of unknown length
*    52480    54484: contig of 2005 bp in length
*    54485    54584: gap of unknown length
*    54585    57588: contig of 3004 bp in length
*    57589    57688: gap of unknown length
*    57689    59983: contig of 2295 bp in length
*    59984    60083: gap of unknown length
*    60084    62833: contig of 2750 bp in length
*    62834    62933: gap of unknown length
*    62934    65460: contig of 2527 bp in length
*    65461    65560: gap of unknown length
*    65561    67597: contig of 2037 bp in length
*    67598    67697: gap of unknown length
*    67698    71010: contig of 3313 bp in length
*    71011    71110: gap of unknown length
*    71111    73099: contig of 1989 bp in length
*    73100    73199: gap of unknown length
*    73200    76659: contig of 3460 bp in length
*    76660    76759: gap of unknown length
*    76760    79328: contig of 2569 bp in length
*    79329    79428: gap of unknown length
*    79429    82950: contig of 3522 bp in length
*    82951    83050: gap of unknown length
*    83051    85337: contig of 2287 bp in length
*    85338    85437: gap of unknown length
*    85438    88632: contig of 3195 bp in length
*    88633    88732: gap of unknown length
*    88733    92607: contig of 3875 bp in length
*    92608    92707: gap of unknown length
*    92708    95823: contig of 3116 bp in length
*    95824    95923: gap of unknown length
*    95924    98417: contig of 2494 bp in length
*    98418    98517: gap of unknown length
*    98518   102227: contig of 3710 bp in length
*   102228   102327: gap of unknown length
*   102328   106214: contig of 3887 bp in length
*   106215   106314: gap of unknown length
*   106315   109534: contig of 3220 bp in length
*   109535   109634: gap of unknown length
*   109635   113077: contig of 3443 bp in length
*   113078   113177: gap of unknown length
*   113178   116869: contig of 3692 bp in length
*   116870   116969: gap of unknown length
*   116970   121793: contig of 4824 bp in length
*   121794   121893: gap of unknown length
*   121894   126597: contig of 4704 bp in length
*   126598   126697: gap of unknown length
*   126698   131888: contig of 5191 bp in length
*   131889   131988: gap of unknown length
*   131989   136453: contig of 4465 bp in length
*   136454   136553: gap of unknown length
*   136554   141386: contig of 4833 bp in length
*   141387   141486: gap of unknown length
*   141487   146790: contig of 5304 bp in length
*   146791   146890: gap of unknown length
*   146891   151544: contig of 4654 bp in length
*   151545   151644: gap of unknown length
*   151645   157304: contig of 5660 bp in length
*   157305   157404: gap of unknown length

  Query Match             26.3%;  Score 827.4;  DB 2;  Length 197893;
  Best Local Similarity   80.9%;  Pred. No. 1.1e-133;
  Matches 1062;  Conservative    0;  Mismatches  216;  Indels   35;  Gaps    7;

Qy       1859 TTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGA 1918
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||| || |
Db     194803 TCTGCAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCCGGAA 194744

Qy       1919 CTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCA 1978
```

```
              | |||||||||||||| || ||||| |||||||| ||||||||||| || ||||||||||
Db     194743 CGGGCTACGACTTGTGCGTACTGGCTATGTGGACGACACAGCCACAACGCTCACGGCCA 194684

Qy       1979 CCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCA 2038
              ||||||| |||||||||||||||||||||||||||||||||||||| || |||||||| |
Db     194683 CCAACATTGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCACAATGCCAGTCTA 194624

Qy       2039 TGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCA 2098
              ||||||||||||| ||||| ||||| ||||| ||||||||||| || ||||| |||||||
Db     194623 TGCACAGCCAGATCCTGGGGGGCACTATGATTCTGGTCATCGGAGGTATCATTGTGGCCA 194564

Qy       2099 CGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCC 2158
              | ||||||||||||||||| |||||||||||||| |||||||||||||||||||| ||||
Db     194563 CACTGCTGGTCTTCATCGTAATCCTCATGGTGCGGTACAAGGTCTGCAACCACGACGCCC 194504

Qy       2159 CCAGCAAGATGGCAGCGGCC---GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGC 2215
              |  |||| |||||||| |||   || |||||||| |||||||||||||| || || ||||
Db     194503 CAGGCAAAATGGCAGCCGCCACTGTCAGCAATGTATACTCGCAGACCAATGGAGCTCAGC 194444

Qy       2216 CACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGG 2275
              | || ||||   |  |    || |  ||    |  || ||||  || || ||||| ||||
Db     194443 CTCCACCTCTGGGTGGGATGCCTGTTGGCCAGCTTCCTCAGGCACCACCCAAGGTCGTGG 194384

Qy       2276 TGCGCAACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCT 2335
              |||| || ||| |  |||||||||   ||||||||||  | ||| ||||||| |||||||
Db     194383 TGCGGAATGAGTTGATGGACTTCAGTACCAGCCTGGCAAGAGCCTGTGACTCTTCTTCCT 194324

Qy       2336 CCAGCTCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCAC 2395
              |||| ||||||||||||||||| ||||| |||||| | ||||  ||||||||| ||||||
Db     194323 CCAGTTCCCTGGGCAGTGGGGAAGCTGCAGGGCTGAGCCGGGGTCCCTGGAGGCTCCCAC 194264

Qy       2396 CCTCCGCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGG 2455
              || | || || || ||||||||||||||||||||| ||||||| || || ||||||||||
Db     194263 CCCCAGCTCCTCGTCCCAAGCCCAGCCTTGACCGCTTGATGGGAGCTTTTGCCTCCCTGG 194204

Qy       2456 ACCTCAAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGG 2515
              ||||||||||||| || || || || ||  | ||||| ||||||||||| || |||||||
Db     194203 ACCTCAAGAGTCAAAGGAAAGAAGAACTTTTAGACTCTAGGACTCCAGCAGGCAGAGGGG 194144

Qy       2516 CTGGGACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGG 2575
              | ||||| || ||| ||||||||||||| |||||||||||||||||||||||||| ||
Db     194143 CAGGGACATCAGCCAGGGGCCACCACTCAGACCGAGAGCCACTGCTGGGGCCCCCAGCCA 194084

Qy       2576 CCCGGGCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCT 2635
              |||| ||||||||||| ||||| |||||  ||||||||||||||||||| |||||||| |
Db     194083 CCCGTGCCAGGAGCCTTCTCCCTTTGCCCCTGGAGGGCAAGGCCAAACGAAGCCACTCTT 194024

Qy       2636 TCGACATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTC 2695
              | ||||||||||||||||| ||||| || || |        | || || ||||||||||
Db     194023 TTGACATGGGGGACTTTGCAGCTGCAGCTGCAG------CTGCTCCTGGTGGCTACAGTC 193970

Qy       2696 CTCCTCGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCT 2755
              | ||| ||  |||||| ||||||||||||||||||||||| |||||||| ||||||||||
Db     193969 CCCCTAGGCGGGTCTCAAACATCTGGACGAAGCGCAGCCTGTCTGTCAATGGCATGCTCT 193910

Qy       2756 TGCCCTTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGG 2815
              ||||||||||||||| |||||||||||| || |||||||| ||||||||||| || ||||
Db     193909 TGCCCTTTGAGGAGAATGACCTGGTGGGAGCTCGGGGGACATTTGGCAGCTCAGAGTGGG 193850

Qy       2816 TGATGGAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGG 2875
              | |||||||| || || |||   |||  ||     |  || | ||      |      |
Db     193849 TAATGGAGAGTACTGTGTAGCCAGGGAAGGAGTACCGTCCCTCCCGCCCTTCCTCAAAGT 193790

Qy       2876 AGAAGGGGAAAGAATC-------TCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGT 2928
               | || |||| || ||        ||||||||||| ||||||||||||||||||||||||
Db     193789 GGGAGAGGAAGGAGTCAGAGTCAGCACTGGCAAGTATTTGTGGAGTTTCCATGGTGATGT 193730

Qy       2929 TTACATCCAGGGACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTC------CCCCCCTAC 2982
              ||||||||||||||||| | |||||||||||||  ||||| ||||      |||  ||||
Db     193729 TTACATCCAGGGACAGTCTTGTCTCCCTGTCAACAGCCTCATGTCCTTCTGCCCAGCTAC 193670

Qy       2983 CCCGCAACACCCACATCACCTCCCCA-----------CCACCCGGCCGGGGTGTGCTCAG 3031
               || |    || |  |||||| ||||            |  ||   ||||||||||||||
Db     193669 ACCCCCTTCCCTATGTCACCTTCCCAGGCCTGTCCCCACCACCCAGCGGGGTGTGCTCAG 193610

Qy       3032 GGAATGTGGACTCGCTCAAATGCCGGACTGAGCCCTGAGTGTTTGGAAAGGCGAGACTCC 3091
              |||||| |||||||||||  || | ||||||||||||||||||||||||||| |    |
Db     193609 GGAATG-GGACTCGCTCATGTGTCAGACTGAGCCCTGAGTGTTTGGAAAGGCTAACAACT 193551

Qy       3092 GCCTTTCTAATCACAAATGTAGCCTACAAGCAAGCGGCTTTGGATTGCTTATG 3144
              ||||||||| |||||||||||| | ||||||||||||||| ||||||||| ||
Db     193550 GCCTTTCTATTCACAAATGTAG-CGACAAGCAAGCGGCTTCGGATTGCTTGTG 193499


RESULT 10
AX179258
LOCUS       AX179258                1887 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 3 from Patent WO0142471.
ACCESSION   AX179258
VERSION     AX179258.1  GI:14598947
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1887)
  AUTHORS   Prayaga,S.K. and Shimkets,R.A.
  TITLE     Polypeptides and polynucleotides encoding same
  JOURNAL   Patent: WO 0142471-A 3 14-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..1887
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Curagen clone AF038458_A"
BASE COUNT      259 a    736 c    577 g    315 t
ORIGIN

  Query Match             20.2%;  Score 634.6;  DB 6;  Length 1887;
  Best Local Similarity   61.3%;  Pred. No. 4.1e-100;
  Matches 1079;  Conservative    0;  Mismatches  664;  Indels   16;  Gaps    3;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
              ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db         63 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 122

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                |    ||||||||       ||  ||  ||  ||  ||  ||  ||     | ||||| ||| |
Db        123 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 182

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
              ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db        183 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 242

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
              |||||| || ||| || | |||| |||||||  ||||||||||| ||||| |     ||
Db        243 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 302

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db        303 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 362

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
               || || |||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db        363 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 422

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
              ||||||||  |   | || |  |  || || || |       |||| ||||||| ||
Db        423 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 482

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
               ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db        483 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 542

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
              | || ||  | || |||| |||||||||||||||| |    |  |||  ||| | ||| |
Db        543 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 602
```

```
Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                 |||||| || ||||||||| |||||  | |||||||| || |||   |   | || ||
Db        603 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 662

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
               || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db        663 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 722

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
              | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db        723 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 782

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db        783 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 842

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
              ||||||||||   ||| | ||||||||||||||| |||||||||||  |  | || || |
Db        843 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 902

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
              | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db        903 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 962

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
              ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db        963 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1022

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                |  ||   | ||    |||||| || |||||  |  |  |||||    |    ||  |
Db       1023 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1082

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
              |||   |||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db       1083 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1142

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db       1143 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1202

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                 ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db       1203 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1262

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
               |      | | ||| |        | || || |  || |  ||| ||||        ||
Db       1263 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1314

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
              ||  | || || ||| |||||| | |       | |    ||  |  |    |||||||||
Db       1315 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1374

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
               | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db       1375 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1434

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                  | ||| || | |   |||||| ||||||      ||||||  ||||  |||| ||||| |||
Db       1435 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1494

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
               |||  || ||||  |||||||   |||||||||||       |||||||||||||  |||
Db       1495 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1554

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
              | |||| |  |||     | || | ||    |   ||||     |   |  |||||||||
Db       1555 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 1614

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
              |||||| | |   | || ||| ||||||| ||| ||   ||||||||||||| || | ||
Db       1615 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 1674

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
               |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db       1675 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 1734

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
              || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db       1735 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 1794

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257
               |||   |   |||   ||
Db       1795 CCCGCCCCGGAGCCCGCGG 1813

RESULT 11
BC003578
LOCUS       BC003578                2549 bp    mRNA    linear   PRI 12-JUL-2001
DEFINITION  Homo sapiens, clone MGC:2656 IMAGE:3537259, mRNA, complete cds.
ACCESSION   BC003578
VERSION     BC003578.1  GI:13097761
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2549)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 8 Row: f Column: 13
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, GenomeScan gene prediction..
FEATURES             Location/Qualifiers
     source          1..2549
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:2656 IMAGE:3537259"
                     /tissue_type="Lung, small cell carcinoma"
                     /clone_lib="NIH_MGC_7"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             453..2339
                     /codon_start=1
                     /product="Unknown (protein for MGC:2656)"
                     /protein_id="AAH03578.1"
                     /db_xref="GI:13097762"
                     /translation="MAILPLLLCLLPLAPASSPPQSATPSPCPRRCRCQTQSLPLSVL
                     CPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANMTGLLHLSLSRNTIRHVAAGA
                     FADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQLAALAAGALDDCAETLED
                     LDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASVPAGAFSRLHKLARLDMTSNRLTT
```

```
                     IPPDPLFSRLPLLARPRGSPASALVLAFGGNPLHCNCELVWLRRLAREDDLEACASPP
                     ALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRAVGDPEPRVRWVSPQG
                     RLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAVELTVGPPPPPQLA
                     NSTSCDPPRDGDPDALTPPSAASASAKVADTGPPTDRGVQVTEHGATAALVQWPDQRP
                     IPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLCVLAVYEDSATGLT
                     ATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFIFVLLMRYKVHG
                     GQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEPAALRAHTVVQLDCEPWGPGHE
                     PVGP"
BASE COUNT      440 a    935 c    738 g    436 t
ORIGIN

  Query Match              20.2%;  Score 634.6;  DB 9;  Length 2549;
  Best Local Similarity    61.3%;  Pred. No. 4e-100;
  Matches 1079;  Conservative    0;  Mismatches  664;  Indels   16;  Gaps    3;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
Db        515 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 574

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
Db        575 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 634

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
Db        635 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 694

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
Db        695 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 754

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
Db        755 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 814

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
Db        815 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 874

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
Db        875 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 934

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
Db        935 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 994

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
Db        995 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 1054

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
Db       1055 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1114

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
Db       1115 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1174

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
Db       1175 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1234

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
Db       1235 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1294

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
Db       1295 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1354

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
Db       1355 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1414

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
Db       1415 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1474

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
Db       1475 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1534

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
Db       1535 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGGGAGGCCACAGCTGCTGTGGAGCT 1594

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
Db       1595 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1654

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
Db       1655 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1714

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
Db       1715 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1766

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
Db       1767 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1826

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
Db       1827 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1886

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
Db       1887 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1946

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
Db       1947 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2006

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
Db       2007 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2066

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
Db       2067 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2126

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
Db       2127 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2186

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
Db       2187 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2246

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257
Db       2247 CCCGCCCCGGAGCCCGCGG 2265


RESULT 12
AX266997
LOCUS       AX266997                1887 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 3 from Patent WO0172827.
ACCESSION   AX266997
VERSION     AX266997.1  GI:16515796
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
  AUTHORS  Glucksmann,M.A.
  TITLE    33395, a novel human leucine-rich repeat family member and use
           thereof
  JOURNAL  Patent: WO 0172827-A 3 04-OCT-2001;
           Millennium Pharmaceuticals, Inc. (US)
FEATURES            Location/Qualifiers
     source         1..1887
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
BASE COUNT     257 a    735 c    579 g    316 t
ORIGIN

  Query Match             20.1%;  Score 633;  DB 6;  Length 1887;
  Best Local Similarity   61.3%;  Pred. No. 7.7e-100;
  Matches 1078;  Conservative    0;  Mismatches  665;  Indels   16;  Gaps    3;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
              ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db         63 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 122

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                |    ||||||||       || || || || || || ||     | ||||| ||| |
Db        123 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 182

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
              ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db        183 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 242

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
              |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db        243 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 302

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db        303 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 362

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
               || || |||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db        363 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 422

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
              ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db        423 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 482

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
               ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db        483 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 542

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
              | || ||  | || |||| |||||||||||||||| |    |  |||  ||| | ||| |
Db        543 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 602

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                 |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db        603 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 662

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
               || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db        663 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 722

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
              | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db        723 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 782

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db        783 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 842

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
              |||||| |||   ||| | |||||||||||||| |||||||||||  |  | || || |
Db        843 CTACTTTTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 902

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
              | |||      ||   || |  |  | |  | ||||  | ||    |||   ||  | ||
Db        903 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 962

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
              ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db        963 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1022

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                |  ||   | ||    |||||| || |||||  |  |  |||||    |    ||  |
Db       1023 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1082

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
              |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db       1083 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1142

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db       1143 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1202

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                 ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db       1203 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1262

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
               |      | | ||| |        | || || |  || |  ||| ||||        ||
Db       1263 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1314

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
              ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db       1315 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1374

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
               | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db       1375 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1434

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                  | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db       1435 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1494

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
               |||  || | ||  ||||||   |||||||||||       |||||||||||||  |||
Db       1495 GTGTATGAGGGCAGTGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1554

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
              | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db       1555 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 1614

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
              |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db       1615 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 1674

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
               |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| ||
Db       1675 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCC 1734

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
              || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db       1735 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 1794

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257
               |||   |   |||   ||
Db       1795 CCCGCCCCGGAGCCCGCGG 1813


RESULT 13
AX266995
LOCUS       AX266995                2558 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 1 from Patent WO0172827.
ACCESSION   AX266995
VERSION     AX266995.1  GI:16515794
KEYWORDS    .
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Glucksmann,M.A.
  TITLE     33395, a novel human leucine-rich repeat family member and use
            thereof
  JOURNAL   Patent: WO 0172827-A 1 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2558
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             435. .2321
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD10239.1"
                     /db_xref="GI:16515795"
                     /translation="MAILPLLLCLLPLAPASSPPQSATPSPCPRRCRCQTQSLPLSVL
                     CPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANMTGLLHLSLSRNTIRHVAAGA
                     FADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQLAALAAGALDDCAETLED
                     LDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASVPAGAFSRLHKLARLDMTSNRLTT
                     IPPDPLFSRLPLLARPRGSPASALVLAFGGNPLHCNCELVWLRRLAREDDLEACASPP
                     ALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRAVGDPEPRVRWVSPQG
                     RLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAVELTVGPPPPPQLA
                     NSTSCDPPRDGDPDALTPPSAASASAKVADTGPPTDRGVQVTEHGATAALVQWPDQRP
                     IPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLCVLAVYEGSATGLT
                     ATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFIFVLLMRYKVHG
                     GQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEPAALRAHTVVRLDCEPWGPGHE
                     PVGP"
BASE COUNT      422 a    956 c    743 g    437 t
ORIGIN

  Query Match            20.1%;  Score 633;  DB 6;  Length 2558;
  Best Local Similarity  61.3%;  Pred. No. 7.5e-100;
  Matches 1078;  Conservative    0;  Mismatches  665;  Indels   16;  Gaps    3;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
Db        497 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 556

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
Db        557 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 616

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
Db        617 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 676

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
Db        677 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 736

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
Db        737 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 796

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
Db        797 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 856

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
Db        857 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 916

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
Db        917 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 976

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
Db        977 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 1036

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
Db       1037 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1096

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
Db       1097 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1156

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
Db       1157 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1216

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
Db       1217 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1276

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
Db       1277 CTACTTTTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1336

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
Db       1337 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1396

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
Db       1397 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1456

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
Db       1457 CCGTGCCCGCGCCTTCCCCAATGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1516

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
Db       1517 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1576

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
Db       1577 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1636

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
Db       1637 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1696

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
Db       1697 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1748

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
Db       1749 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1808

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
Db       1809 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1868

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
Db       1869 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1928

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
Db       1929 GTGTATGAGGGCAGTGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1988

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
Db       1989 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2048

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
Db       2049 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2108
```

```
Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCAG------CAAGATGGCAGCGGCC 2178
               |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| ||
Db       2109 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCC 2168

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
              || ||||  || | ||| ||||||||||||||| |   ||  | || |    ||| ||
Db       2169 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2228

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257
               |||   |   |||   ||
Db       2229 CCCGCCCCGGAGCCCGCGG 2247


RESULT 14
BC015581
LOCUS       BC015581                2458 bp    mRNA    linear   PRI 17-JAN-2002
DEFINITION  Homo sapiens, clone MGC:23278 IMAGE:4637053, mRNA, complete cds.
ACCESSION   BC015581
VERSION     BC015581.1  GI:18182830
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2458)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site:       http://www.nisc.nih.gov/
            Contact:        nisc_mgc@nhgri.nih.gov
            Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
            Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 33 Row: c Column: 15
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1..2458
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MGC:23278 IMAGE:4637053"
                     /tissue_type="Lung, large cell carcinoma"
                     /clone_lib="NIH_MGC_18"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             1186..2160
                     /codon_start=1
                     /product="Unknown (protein for MGC:23278)"
                     /protein_id="AAH15581.1"
                     /db_xref="GI:18182831"
                     /translation="MHWVGPDDRLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATN
                     PAGEATARVELRVLALPHGGNSSAEGGRPGPSDIAASARTAAEGEGTLESEPAVQVTE
                     VTATSGLVSWGPGRPADPVWMFQIQYNSSEDETLIYRIVPASSHHFLLKHLVPGADYD
                     LCLLALSPAAGPSDLTATRLLGCAHFSTLPASPLCHALQAHVLGGTLTVAVGGVLVAA
                     LLVFTVALLVRGRGAGNGRLPLKLSHVQSQTNGGPSPTPKAHPPRSPPPRPQRSCSLD
                     LGDAGCYGYARRLGGAWARRSHSVHGGLLGAGCRGVGGSAERLEESVV"
BASE COUNT      358 a    885 c    783 g    432 t
ORIGIN

  Query Match            19.8%;  Score 621;  DB 9;  Length 2458;
  Best Local Similarity  62.2%;  Pred. No. 9e-98;
  Matches 1088;  Conservative    0;  Mismatches  610;  Indels   51;  Gaps    5;

Qy        520 GACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGC 579
              | |||||||||   |  ||| ||||||||||| ||||| ||||| ||  |  ||||| ||
Db        295 GCCGCCTGCCCGCTGCCCTGCGTCTGCCAGAACCTGTCCGAGTCGCTCAGCACCCTCTGT 354

Qy        580 CCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGC 639
               ||  |   || ||||| ||||| || ||  |  | ||||||||| ||||||||||||||
Db        355 GCCCACCGAGGCCTGCTGTTTGTGCCGCCCAACGTGGACCGGCGCACAGTGGAGCTGCGG 414

Qy        640 CTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTG 699
              ||||  | ||||||||||     | |  | | ||  |||||   |||||||||||| |||
Db        415 CTGGCTGACAACTTCATCCAGGCCCTGGGGCCCCCTGACTTCCGCAACATGACGGGACTG 474

Qy        700 GTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGAC 759
              ||||||||||| |||||  | ||  |||||| || |||   | ||    |||||  ||||
Db        475 GTGGACCTGACACTGTCTCGCAATGCCATCACCCGCATTGGGGCCCGCGCCTTTGGGGAC 534

Qy        760 CTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGAC 819
              ||||||||||| || ||||| || |||||| ||||  |||||      || ||    |
Db        535 CTCGAGAGCCTGCGTTCCCTCCACCTTGACGGCAACAGGCTGGTGGAGCTGGGCACCGGG 594

Qy        820 ACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGC 879
              | |||||||||||  ||||| ||||||||||| ||| | | |  ||||||||||||| ||
Db        595 AGCCTCCGGGGCCCCGTCAATCTGCAGCACCTCATCCTCAGCGGCAACCAGCTGGGCCGC 654

Qy        880 ATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTAC 939
              |||||    |  || || || ||||||||   ||   ||||||| ||||||||| ||||||
Db        655 ATCGCGCCGGGAGCCTTCGACGACTTCCTAGAGAGCCTGGAGGACCTGGACCTGTCCTAC 714

Qy        940 AACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTG 999
              ||||||||||      |||| |||| |  | |  |   ||||     ||| |||   ||
Db        715 AACAACCTCCGGCAGGTGCCCTGGGCCGGCATCGGCGCCATGCCTGCCCTGCACACCCTC 774

Qy       1000 AGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAA 1059
              | ||||||||| |||||  | ||     |  ||   ||| |||| ||  | ||     |
Db        775 AACCTGGACCATAACCTTATTGACGCACTGCCCCCAGGCGCCTTCGCCCAGCTCGGTCAG 834

Qy       1060 CTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTT 1119
              ||  |||||||||| ||||||||||| || |||   | ||| | || || ||  | ||
Db        835 CTCTCCCGCCTGGACCTCACCTCCAACCGCCTGGCCACGCTGGCTCCGGACCCGCTTTTC 894

Qy       1120 GCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGT 1179
               | ||    |  | |            ||     ||| | || | ||| ||    | ||
Db        895 TCTCGTGGGCGTGAT------------GCAGAGGCCTCTCCCGCCCCCTGGTGCTGAGC 942

Qy       1180 TTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGG 1239
              ||| | || ||||| || |||||||| |||||||| || ||||||||| |||| | ||||
Db        943 TTTAGCGGGAACCCCCTGCACTGCAACTGTGAGCTGCTGTGGCTGCGGCGGCTGGCGCGG 1002

Qy       1240 GACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCAT 1299
                 || |||||||||||| || | ||||||    |||||     || ||||||||||||
Db       1003 CCGGACGACCTGGAAACGTGCGCCTCCCCGCCCGGCCTGGCCGGCCGCTACTTCTGGGCA 1062

Qy       1300 GTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTG 1359
              ||||  ||||  |||||    || |||||||| |||||  |||  ||||| ||     |
Db       1063 GTGCCCGAGGGCGAGTTCTCCTGTGAGCCGCCCCTCATTGCCCGCCACACGCAGCGCCTC 1122

Qy       1360 CTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCC 1419
                ||| |||||| ||||||  |||||| ||   ||||   ||| | || ||||||   ||
Db       1123 TGGGTGCTGGAAGGCCAGCGGGCCACGCTGCGGTGCCGGGCCCTGGGTGACCCCGCGCCT 1182

Qy       1420 CTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTC 1479
                 || |||||||| |  || || |||||  |||| || ||||||||  |  ||   |
Db       1183 ACCATGCACTGGGTCGGTCCTGACGACCGGTTGGTTGGCAACTCCTCCCGAGCCCGGGCT 1242
```

```
Qy       1480 TATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 1539
              |    ||| || ||| | || ||     | |||    ||  ||||  ||| | || ||||
Db       1243 TTCCCCAACGGGACCTTAGAGATTGGGGTGACCGGCGCTGGGGACGCTGGGGGCTACACC 1302

Qy       1540 TGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAG 1599
              ||||| ||  ||||     | || |||||||| ||    || ||  |     |  |    
Db       1303 TGCATCGCCACCAACCCTGCTGGTGAGGCCACAGCCCGAGTAGAACTGCGGGTGCTGGCC 1362

Qy       1600 CTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGAC 1659
               |||| ||     | |||||||    |      |  | ||      || | | ||| |||
Db       1363 TTGCCCCATGGTGGGAACAGCAGTGCCGAGGGGGGCCGCC------CCGGGCCCTCGGAC 1416

Qy       1660 ATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCC 1719
              ||| | | ||||      ||       |  ||||  || |    ||     |||||    
Db       1417 ATCGCCGCCTCCGCTCGCACTGCTGCCGAGGGTGAGGGGACGCTGGAGTCTGAGCC---- 1472

Qy       1720 AAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTC 1779
                                || ||||  ||| | || |||||| ||||||| |  ||||| 
Db       1473 -----------------AGCCGTGCAGGTGACGGAGGTGACCGCCACCTCAGGGCTGGTG 1515

Qy       1780 AAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGC 1839
              |  |||  |  | |   | ||||   ||  |||  ||||| |||  | ||||||||| ||
Db       1516 AGCTGGGGTCCCGGGCGGCCAGCCGACCCAGTGTGGATGTTCCAAATCCAGTACAACAGC 1575

Qy       1840 TCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTC 1899
                 || ||||||   || || ||| ||||  |||||||||||| | |   |||| || | 
Db       1576 AGCGAAGATGAGACCCTCATCTACCGGATTGTCCCAGCCTCCAGCCACCACTTCCTGCTG 1635

Qy       1900 AACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACA 1959
              ||  |||| ||  | | | ||| ||| ||| | ||  |||||||| |||     | | | 
Db       1636 AAGCACCTCGTCCCCGGCGCTGACTATGACCTCTGCCTGCTGGCCTTGTCACCGGCCGCT 1695

Qy       1960 G---CCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCT 2016
              |   || |    |||||||||||||   |  ||||||| ||||| |||| |||   | | 
Db       1696 GGGCCCTCTGACCTCACGGCCACCAGGCTGCTGGGCTGTGCCCATTTCTCCACGCTGCCG 1755

Qy       2017 GACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTC 2076
              | ||  || | ||||||  || ||||   |||  | ||||||| ||| ||| || | ||| |
Db       1756 GCCTCGCCCCTGTGCCACGCCCTGCAGGCCCACGTGCTGGGCGGGACCCTGACCGTGGCC 1815

Qy       2077 ATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTAC 2136
               | |||||  |  | |||||  |  | ||||||||  ||   | |  ||||| ||   |
Db       1816 GTGGGGGGTGTGCTGGTGGCTGCCTTACTGGTCTTCACTGTGGCCTTGCTGGTTCGGGGC 1875

Qy       2137 AAGGTCTGCAACCACGAGGCCCCCAGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCG 2196
                ||    |    | |    || |  |   | |          | ||| | ||  | || 
Db       1876 CGGGGGGCCGGAAATGGCCGCCTCCCCCTCAAG---------CTCAGCCACGTCCAGTCC 1926

Qy       2197 CAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAG 2256
              |||||||| || | ||     ||  | || |    |   ||  |  ||  ||| ||||  
Db       1927 CAGACCAATGGAGGCCCCAGCCCCACACCCAAGGCCCACCCGCCGCGGAGCCCCCCGCCC 1986

Qy       2257 GGCCCGCCG 2265
               | |||| |
Db       1987 CGGCCGCAG 1995


RESULT 15
AX358336
LOCUS       AX358336                1653 bp    DNA     linear   PAT 13-FEB-2002
DEFINITION  Sequence 7 from Patent WO0202604.
ACCESSION   AX358336
VERSION     AX358336.1  GI:18674972
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Thayer,E.C., Sheppard,P.O. and Presnell,S.R.
  TITLE     Leucine-rich repeat proteins, zlrr7, zlrr8 and zlrr9
  JOURNAL   Patent: WO 0202604-A 7 10-JAN-2002;
            ZymoGenetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1653
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1..>1653
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD23323.1"
                     /db_xref="GI:18674973"
                     /translation="MAPPLLLLLLASGAAACPLPCVCQNLSESLSTLCAHRGLLFVPP
                     NVDRRTVELRLADNFIQALGPPDFRNMTGLVDLTLSRNAITRIGARAFGDLESLRSLH
                     LDGNRLVELGTGSLRGPVNLQHLILSGNQLGRIAPGAFDDFLESLEDLDLSYNNLRQV
                     PWAGIGAMPALHTLNLDHNLIDALPPGAFAQLGQLSRLDLTSNRLATLAPDPLFSRGR
                     DAEASPAPLVLSFSGNPLHCNCELLWLRRLARPDDLETCASPPGLAGRYFWAVPEGEF
                     SCEPPLIARHTQRLWVLEGQRATLRCRALGDPAPTMHWVGPDDRLVGNSSRARAFPNG
                     TLEIGVTGAGDAGGYTCIATNPAGEATARVELRVLALPHGGNSSAEGGRPGPSDIAAS
                     ARTAAEGEGTLESEPAVQVTEVTATSGLVSWGPGRPADPVWMFQIQYNSSEDETLIYR
                     IVPASSHHFLLKHLVPGADYDLCLLALSPAAGPSDLTATRLLGCAHFSTLPASPLCHA
                     LQAHVLGGTLTVAVGGVLVAALTGLHCGPCWFGAGGPEMAASP"
BASE COUNT      223 a    617 c    526 g    287 t
ORIGIN

  Query Match             19.3%;  Score 608;  DB 6;  Length 1653;
  Best Local Similarity   63.7%;  Pred. No. 1.7e-95;
  Matches 1025;  Conservative    0;  Mismatches  540;  Indels   43;  Gaps    5;

Qy        520 GACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGC 579
              | |||||||||   |  ||| ||||||||||| ||||| | |||| ||  |  ||||| || 
Db         43 GCCGCCTGCCCGCTGCCCTGCGTCTGCCAGAACCTGTCCGAGTCGCTCAGCACCCTCTGT 102

Qy        580 CCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGC 639
               ||  |   || ||||| ||||| || ||  |  | |||||||| ||||||||||||||| 
Db        103 GCCCACCGAGGCCTGCTGTTTGTGCCGCCCAACGTGGACCGGCGCACAGTGGAGCTGCGG 162

Qy        640 CTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTG 699
              ||||  | ||||||||||     | |  | | ||  |||||   |||||||||||| |||
Db        163 CTGGCTGACAACTTCATCCAGGCCCTGGGGCCCCCTGACTTCCGCAACATGACGGGACTG 222

Qy        700 GTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGAC 759
              ||||||||||| |||||  | ||  ||||||| || |||   | ||    |||||  ||||
Db        223 GTGGACCTGACACTGTCTCGCAATGCCATCACCCGCATTGGGGCCCGCGCCTTTGGGGAC 282

Qy        760 CTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGAC 819
              ||||||||||| || ||||| || |||||| ||||  |||||      || ||    |  
Db        283 CTCGAGAGCCTGCGTTCCCTCCACCTTGACGGCAACAGGCTGGTGGAGCTGGGCACCGGG 342

Qy        820 ACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGC 879
              | ||||||||||||  ||||| ||||||||||| ||| | | |  ||||||||||||| ||
Db        343 AGCCTCCGGGGCCCCGTCAATCTGCAGCACCTCATCCTCAGCGGCAACCAGCTGGGCCGC 402

Qy        880 ATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTAC 939
              |||||    |  || |  || |||||||||   ||   ||||||| ||||||||| ||||||
Db        403 ATCGCGCCGGGAGCCTTCGACGACTTCCTAGAGAGCCTGGAGGACCTGGACCTGTCCTAC 462

Qy        940 AACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTG 999
              ||||||||||      |||| |||| |  | |  |   ||||     ||| |||   || 
Db        463 AACAACCTCCGGCAGGTGCCCTGGGCCGGCATCGGCGCCATGCCTGCCCTGCACACCCTC 522

Qy       1000 AGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAA 1059
              | ||||||||| |||||  | ||     |  ||   ||| |||| ||  | ||     | 
Db        523 AACCTGGACCATAACCTTATTGACGCACTGCCCCCAGGCGCCTTCGCCCAGCTCGGTCAG 582

Qy       1060 CTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTT 1119
              ||  |||||||||| |||||||||| || |||   | |||| | || || ||  | || 
Db        583 CTCTCCCGCCTGGACCTCACCTCCAACCGCCTGGCCACGCTGGCTCCGGACCCGCTTTTC 642

Qy       1120 GCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGT 1179
               | ||    |  | |            ||     ||| | || | ||| ||    | || 
Db        643 TCTCGTGGGCGTGAT------------GCAGAGGCCTCTCCCGCCCCCCTGGTGCTGAGC 690

Qy       1180 TTTGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGG 1239
              ||| | || ||||| || |||||||| |||||||| || ||||||||| |||| | ||||
```

```
Db        691 TTTAGCGGGAACCCCCTGCACTGCAACTGTGAGCTGCTGTGGCTGCGGCGGCTGGCGCGG 750

Qy       1240 GACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCAT 1299
                 || ||||||||||| || | ||||||    |||||    || ||||||||||||
Db        751 CCGGACGACCTGGAAACGTGCGCCTCCCCGCCCGGCCTGGCCGGCCGCTACTTCTGGGCA 810

Qy       1300 GTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTG 1359
              ||||  ||||  |||||    || |||||||| |||||  |||  ||||| ||     |
Db        811 GTGCCCGAGGGCGAGTTCTCCTGTGAGCCGCCCCTCATTGCCCGCCACACGCAGCGCCTC 870

Qy       1360 CTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCC 1419
                ||| ||||| ||||||  |||||| ||   ||||   ||| | || ||||||   ||
Db        871 TGGGTGCTGGAAGGCCAGCGGGCCACGCTGCGGTGCCGGGCCCTGGGTGACCCCGCGCCT 930

Qy       1420 CTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTC 1479
                 || ||||||||| |  || || |||||  |||| || ||||||||  |  ||   |
Db        931 ACCATGCACTGGGTCGGTCCTGACGACCGGTTGGTTGGCAACTCCTCCCGAGCCCGGGCT 990

Qy       1480 TATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 1539
              |    ||| || ||| | || ||     | |||    ||  ||||  ||| | || ||||
Db        991 TTCCCCAACGGGACCTTAGAGATTGGGGTGACCGGCGCTGGGGACGCTGGGGGCTACACC 1050

Qy       1540 TGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAG 1599
              ||||| ||  ||||  |||| || |||||||| |||   || ||  |     |  |
Db       1051 TGCATCGCCACCAACCCTGCTGGTGAGGCCACAGCCCGAGTAGAACTGCGGGTGCTGGCC 1110

Qy       1600 CTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGAC 1659
               |||| ||     | |||||||    |      |  | ||      || | | ||| |||
Db       1111 TTGCCCCATGGTGGGAACAGCAGTGCCGAGGGGGGCCGCC------CCGGGCCCTCGGAC 1164

Qy       1660 ATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCC 1719
              ||| | | ||||      ||       |  ||||  || |    ||     |||||
Db       1165 ATCGCCGCCTCCGCTCGCACTGCTGCCGAGGGTGAGGGGACGCTGGAGTCTGAGCC---- 1220

Qy       1720 AAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTC 1779
                                || ||||  ||| | || |||||| ||||||| |  |||||
Db       1221 -----------------AGCCGTGCAGGTGACGGAGGTGACCGCCACCTCAGGGCTGGTG 1263

Qy       1780 AAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGC 1839
              |  |||  |  | |   | ||||   ||  |||  ||||| |||  | ||||||||| ||
Db       1264 AGCTGGGGTCCCGGGCGGCCAGCCGACCCAGTGTGGATGTTCCAAATCCAGTACAACAGC 1323

Qy       1840 TCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTC 1899
                 || ||||||   || || ||| ||||  ||||||||||||| | |   |||| || |
Db       1324 AGCGAAGATGAGACCCTCATCTACCGGATTGTCCCAGCCTCCAGCCACCACTTCCTGCTG 1383

Qy       1900 AACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACA 1959
              ||  |||| ||  | ||  ||| ||| ||| | ||  |||||||| |||     | | |
Db       1384 AAGCACCTCGTCCCCGGCGCTGACTATGACCTCTGCCTGCTGGCCTTGTCACCGGCCGCT 1443

Qy       1960 G---CCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCT 2016
              |   || |    |||||||||||||   |  ||||||| ||||| |||| |||   | |
Db       1444 GGGCCCTCTGACCTCACGGCCACCAGGCTGCTGGGCTGTGCCCATTTCTCCACGCTGCCG 1503

Qy       2017 GACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTC 2076
              | ||  || | ||||||  || ||||   |||  | ||||||||| ||| ||| | ||| |
Db       1504 GCCTCGCCCCTGTGCCACGCCCTGCAGGCCCACGTGCTGGGCGGGACCCTGACCGTGGCC 1563

Qy       2077 ATCGGGGCATCATCGTGGCCACGC-TGCTGGTCTTCATCGTCATCCT 2123
               | ||||  |  | |||||  ||| | ||||||||||  ||   |||
Db       1564 GTGGGGGTGTGCTGGTGGCTGCGCTTACTGGTCTTCACTGTGGCCCT 1611

Search completed: June  8, 2003, 18:21:50
Job time : 9339 secs
```

```
                          GenCore version 5.1.6
                 Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          June  8, 2003, 15:09:17 ; Search time 664 Seconds
                                            (without alignments)
                                            10663.076 Million cell updates/sec

Title:           US-09-831-846-1
Perfect score:   3144
Sequence:        1 gcctggctccctctcgctga..........gcggctttggattgcttatg 3144

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_101002:*
                 1:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                 2:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                 3:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                 4:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                 5:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                 6:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                 7:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                 8:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                 9:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                 10:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                 11:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                 12:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                 13:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                 14:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                 15:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                 16:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                 17:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                 18:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                 19:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                 20:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                 21:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                 22:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                 23:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                 24:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                       %
Result               Query
  No.     Score    Match Length DB  ID                  Description
  --------------------------------------------------------------------------
     1     3144    100.0   3144  21  AAA40083             Human brain-specif
     2     3144    100.0   3144  21  AAA60605             Human hh00149 prot
     3   2780.2     88.4   2818  22  AAS44992             cDNA encoding nove
     4   2780.2     88.4   2818  22  AAI58215             Human polynucleoti
     5      969     30.8    969  23  AAS77723             DNA encoding novel
     6    678.8     21.6    771  22  AAI60001             Human polynucleoti
     7    634.6     20.2   1887  22  AAH25568             Nucleotide sequenc
     8    634.6     20.2   1887  22  AAH78215             Nucleotide sequenc
     9      633     20.1   1887  22  AAI66986             Human LRR family m
    10      633     20.1   2637  22  AAI66985             Human leucine-rich
    11    622.6     19.8   2369  24  AAD38695             Human LP220 secret
    12    621.8     19.8   2855  24  ABK70010             cDNA encoding huma
    13    618.8     19.7   2316  22  AAH78204             Nucleotide sequenc
    14      608     19.3   1653  24  AAD28124             Human leucine-rich
    15    554.6     17.6   1653  24  AAD28125             Human leucine-rich
    16      546     17.4   3012  24  ABN99416             Human secreted pro
    17    544.4     17.3   2301  22  AAH78197             Nucleotide sequenc
    18    477.2     15.2   1440  22  AAH78214             Nucleotide sequenc
    19    477.2     15.2   8269  22  AAI99642             Human expressed po
    20    477.2     15.2   8269  22  ABA06793             Human genomic DNA
    21    477.2     15.2   8269  22  AAK90941             Human digestive sy
    22    477.2     15.2   8269  22  AAI64010             Human polynucleoti
    23    477.2     15.2   8269  22  AAS31472             Human DNA for a no
    24    477.2     15.2   8269  22  AAS31976             Human liver associ
    25    477.2     15.2   8269  22  AAS33436             DNA encoding human
    26    477.2     15.2   8269  22  AAS35069             DNA #19 encoding h
    27    477.2     15.2   8269  24  ABQ66796             Human polynucleoti
    28    477.2     15.2   8269  24  ABN90331             Human liver antige
    29      471     15.0   1510  23  AAS70133             DNA encoding novel
    30      471     15.0   1602  22  AAH78203             Nucleotide sequenc
    31    462.2     14.7   1582  23  AAS69014             DNA encoding novel
    32    461.4     14.7   1872  22  AAF76853             Human secreted pro
    33      396     12.6    441  22  AAS45180             cDNA encoding nove
    34    298.2      9.5   1268  22  AAS29522             Human endocrine po
    35    226.2      7.2    541  22  AAI99624             Human expressed po
    36    226.2      7.2    541  22  AAS33219             DNA encoding human
    37    204.2      6.5    432  24  ABK33084             DNA encoding novel
    38    204.2      6.5    432  24  AAD28122             Human leucine-rich
    39      194      6.2    902  22  ABA06401             Human cDNA SEQ ID
    40      194      6.2    902  22  AAS29645             Human endocrine po
    41      194      6.2    902  22  AAI63902             Human polynucleoti
    42      194      6.2    902  22  AAS31239             Human cDNA encodin
    43      194      6.2    902  22  AAS34784             cDNA encoding nove
    44      194      6.2    902  24  ABQ66563             Human polynucleoti
    45    193.2      6.1    902  24  ABL90356             Human polynucleoti

                                 ALIGNMENTS

RESULT 1
AAA40083
ID   AAA40083 standard; cDNA; 3144 BP.
XX
AC   AAA40083;
XX
DT   19-OCT-2000  (first entry)
XX
DE   Human brain-specific transmembrane glycoprotein encoding cDNA.
XX
KW   Transmembrane glycoprotein; human; brain; cellular signal transducer;
KW   neuroactive; neuroprotective; cerebroprotective; drug development;
KW   treatment; nervous disease; diagnostic; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             466..2835
FT                   /*tag=  a
FT                   /product= "transmembrane glycoprotein"
XX
PN   WO200031256-A1.
XX
PD   02-JUN-2000.
XX
PF   18-NOV-1999;   99WO-JP06449.
XX
PR   20-NOV-1998;   98JP-0331727.
XX
PA   (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
PA   (KAZU-) KAZUSA DNA RES INST.
XX
```

```
PI   Funahashi S,  Miyata S,  Nomura N,  Nagase T,  Ohara O;
XX
DR   WPI; 2000-411954/35.
DR   P-PSDB; AAB09968.
XX
PT   Gene encoding a brain-specific transmembrane glycoprotein that has a
PT   typical PDZ protein binding motif and functions as a cellular signal
PT   transducer, useful in developing drugs for treating nervous diseases -
XX
PS   Claim 1b; Page 43-51; 63pp; Japanese.
XX
CC   This invention describes a novel gene encoding a human brain-specific
CC   transmembrane glycoprotein that has a typical PDZ protein binding motif
CC   and functions as a cellular signal transducer. The product of the
CC   invention has neuroactive, neuroprotective and cerebroprotective
CC   activity. The gene and encoded protein are useful in developing drugs
CC   for treating nervous diseases, and also for studying functions of the
CC   nervous system or onset mechanism of nerve-related diseases. They may
CC   also be used for the production of diagnostic reagents such as
CC   oligonucleotide probes and antibodies for detecting proteins comprising
CC   PDZ domains and the nucleic acids that encode them. They may also be
CC   used to produce and identify modulators of the proteins expression and
CC   activity such as antisense sequences and antibodies. This sequence
CC   encodes the human brain-specific transmembrane glycoprotein described in
CC   the method of the invention.
XX
SQ   Sequence 3144 BP; 589 A; 1027 C; 919 G; 609 T; 0 other;

  Query Match             100.0%;  Score 3144;  DB 21;  Length 3144;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3144;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACACATACACAACCCGGCAGGCTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACACATACACAACCCGGCAGGCTCG 60

Qy         61 TCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGT 120

Qy        121 TCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTT 180

Qy        181 CTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGC 240

Qy        241 CATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCA 300

Qy        301 TCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAA 360

Qy        361 GGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCA 420

Qy        421 GCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTT 480

Qy        481 GGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGT 540

Qy        541 GTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTT 600

Qy        601 GTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATC 660

Qy        661 CACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGG 720

Qy        721 AACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTG 780

Qy        781 CATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAAC 840

Qy        841 CTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAG 900

Qy        901 GACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCG 960

Qy        961 TGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTG 1020

Qy       1021 GATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACC 1080

Qy       1081 TCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCT 1140

Qy       1141 TTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCAC 1200

Qy       1201 TGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGT 1260

Qy       1261 GGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTG 1320

Qy       1321 TGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCG 1380

Qy       1381 GCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCC 1440

Qy       1441 GATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGAC 1500

Qy       1501 ATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCC 1560

Qy       1561 GGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGC 1620

Qy       1621 ACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACC 1680

Qy       1681 AGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1681 AGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCT 1740

Qy       1741 GTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCA 1800

Qy       1801 GCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATT 1860

Qy       1861 TACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACT 1920

Qy       1921 GGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACC 1980

Qy       1981 AACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATG 2040

Qy       2041 CACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACG 2100

Qy       2101 CTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCC 2160

Qy       2161 AGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCG 2220

Qy       2221 CCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGC 2280

Qy       2281 AACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGC 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGC 2340

Qy       2341 TCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACCCTCC 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACCCTCC 2400

Qy       2401 GCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTC 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTC 2460

Qy       2461 AAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGG 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGG 2520

Qy       2521 ACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGG 2580

Qy       2581 GCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGAC 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGAC 2640

Qy       2641 ATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 ATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCT 2700

Qy       2701 CGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCC 2760

Qy       2761 TTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATG 2820

Qy       2821 GAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGGAGAAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGGAGAAG 2880

Qy       2881 GGGAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGG 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGG 2940

Qy       2941 ACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCA 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCA 3000

Qy       3001 CCTCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACT 3060

Qy       3061 GAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAA 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAA 3120

Qy       3121 GCAAGCGGCTTTGGATTGCTTATG 3144
              ||||||||||||||||||||||||
Db       3121 GCAAGCGGCTTTGGATTGCTTATG 3144


RESULT 2
AAA60605
ID   AAA60605 standard; cDNA; 3144 BP.
XX
AC   AAA60605;
XX
DT   19-OCT-2000  (first entry)
XX
DE   Human hh00149 protein encoding cDNA SEQ ID NO:3.
XX
KW   Human; ubiquitin-like protein; 149Y2H#151; hh00149; brain;
KW   2-hybrid screening; neuroprotective; signal transducer;
KW   nervous system disease; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             466..2835
FT                   /*tag=  a
FT                   /product= "hh00149"
XX
PN   WO200031255-A1.
XX
PD   02-JUN-2000.
XX
PF   18-NOV-1999;   99WO-JP06448.
XX
PR   20-NOV-1998;   98JP-0331701.
XX
PA   (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI   Funahashi S,  Miyata S;
XX
DR   WPI; 2000-400066/34.
DR   P-PSDB; AAB12448.
XX
PT   Gene encoding ubiquitin-like protein which interacts with protein
PT   hh00149, useful in the diagnosis and treatment of diseases associated
PT   with the nervous system -
XX
PS   Example 2; Page 64-73; 88pp; Japanese.
XX
CC   The present invention describes a ubiquitin-like protein, designated
CC   149Y2H#151, which interacts with protein hh00149 expressed specifically
CC   in the brain. The 149Y2H#151 protein has neuroprotective activity, and
CC   is a signal transducer. The 149Y2H#151 gene and encoded protein are
CC   useful in the diagnosis and treatment of diseases associated with the
```

```
CC   nervous system. The protein can interact with protein hh00149 expressed
CC   specifically in the brain. The present sequence encodes the human
CC   hh00149 protein.
XX
SQ   Sequence 3144 BP; 589 A; 1027 C; 919 G; 609 T; 0 other;

  Query Match             100.0%;  Score 3144;  DB 21;  Length 3144;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3144;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACACATACACAACCCGGCAGGCTCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCTGGCTCCCTCTCGCTGAGACACACATACACTCACACATACACAACCCGGCAGGCTCG 60

Qy         61 TCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGT 120

Qy        121 TCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTT 180

Qy        181 CTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGC 240

Qy        241 CATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCA 300

Qy        301 TCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAA 360

Qy        361 GGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCA 420

Qy        421 GCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTT 480

Qy        481 GGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGT 540

Qy        541 GTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTT 600

Qy        601 GTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATC 660

Qy        661 CACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGG 720

Qy        721 AACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTG 780

Qy        781 CATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAAC 840

Qy        841 CTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAG 900

Qy        901 GACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCG 960

Qy        961 TGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTG 1020

Qy       1021 GATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACC 1080

Qy       1081 TCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCT 1140

Qy       1141 TTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCAC 1200

Qy       1201 TGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGT 1260

Qy       1261 GGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTG 1320

Qy       1321 TGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCG 1380

Qy       1381 GCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCC 1440

Qy       1441 GATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGAC 1500

Qy       1501 ATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCC 1560

Qy       1561 GGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGC 1620

Qy       1621 ACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACC 1680

Qy       1681 AGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCT 1740

Qy       1741 GTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCA 1800

Qy       1801 GCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATT 1860

Qy       1861 TACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACT 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACT 1920

Qy       1921 GGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACC 1980

Qy       1981 AACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATG 2040
```

```
Qy       2041 CACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACG 2100

Qy       2101 CTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCC 2160

Qy       2161 AGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCG 2220

Qy       2221 CCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGC 2280

Qy       2281 AACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGC 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGC 2340

Qy       2341 TCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACCCTCC 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACCCTCC 2400

Qy       2401 GCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTC 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTC 2460

Qy       2461 AAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGG 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGG 2520

Qy       2521 ACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGG 2580

Qy       2581 GCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGAC 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGAC 2640

Qy       2641 ATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 ATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCT 2700

Qy       2701 CGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCC 2760

Qy       2761 TTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATG 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATG 2820

Qy       2821 GAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGAGAAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGAGAAG 2880

Qy       2881 GGGAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGG 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGG 2940

Qy       2941 ACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCA 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCA 3000

Qy       3001 CCTCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACT 3060

Qy       3061 GAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAA 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAA 3120

Qy       3121 GCAAGCGGCTTTGGATTGCTTATG 3144
              ||||||||||||||||||||||||
Db       3121 GCAAGCGGCTTTGGATTGCTTATG 3144


RESULT 3
AAS44992
ID   AAS44992 standard; cDNA; 2818 BP.
XX
AC   AAS44992;
XX
DT   18-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human secretory protein, Seq ID No 73.
XX
KW   Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW   ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW   transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW   amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW   ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW   gut protection; lung; liver fibrosis; immune deficiency; infection;
KW   severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW   multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW   fertility; analgesic; pain; antigen; ss.
XX
OS   Homo sapiens.
XX
PN   WO200166689-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US04942.
XX
PR   07-MAR-2000; 2000US-0519705.
PR   19-MAY-2000; 2000US-0574454.
PR   17-JUN-2000; 2000US-0596193.
PR   14-JUL-2000; 2000US-0616847.
PR   19-SEP-2000; 2000US-0665363.
PR   20-OCT-2000; 2000US-0693267.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Xu C,  Wehrman T,  Ren F,  Ma Y,  Zhou P;
PI   Zhao QA,  Yang Y,  Drmanac RT,  Zhang J,  Chen R,  Xue AJ,  Wang J;
XX
DR   WPI; 2001-589934/66.
DR   P-PSDB; AAU28092.
XX
PT   Novel polypeptides and nucleic acids obtained from cDNA libraries
PT   prepared from various human tissues, for diagnosis and treatment of
PT   cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS   Claim 1; SEQ ID No 73; 107pp; English.
XX
CC   The invention relates to novel isolated human secreted polypeptides (I)
CC   and polynucleotides (II). (I) and (II) are useful for treating
CC   inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC   ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC   involved in increasing haematopoiesis, stem cell survival, bone growth
CC   and remodeling. (I), (II) and modulators of (II) are useful for
CC   prophylaxis or treatment of one or more cancers. (II) is also useful for
CC   creating transgenic animals useful for studying the in vivo activities of
CC   the polypeptide as well as for studying modulators of the polypeptides.
CC   (I) induces the proliferation of neural cells and regeneration of nerve
CC   and brain tissue and is useful for the treatment of central and
CC   peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC   Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC   sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC   activity, regulation of haematopoiesis and is useful for treating myeloid
CC   or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC   and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC   tissue growth, and in tissue repair, healing of burns, incisions,
CC   ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC   disorders, or periodontal disease. Furthermore, (I) is also useful for
```

```
CC   gut protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues, various immune deficiencies and
CC   disorders including severe combined immunodeficiency (SCID), bacterial or
CC   fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC   rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC   reactions and conditions, such as asthma or other respiratory problems.
CC   In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC   fertility, metabolism, catabolism, anabolism, storage or elimination of
CC   dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC   analgesic effects or other pain reducing effects, immunoglobulin like
CC   activity and can act as an antigen in a vaccine composition to raise an
CC   immune response. AAS44920-AAS45295 represent novel human secreted protein
CC   coding sequences of the invention.
XX
SQ   Sequence 2818 BP; 523 A; 933 C; 825 G; 537 T; 0 other;

  Query Match             88.4%;  Score 2780.2;  DB 22;  Length 2818;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2785;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         43 CACAACCCGGCAGGCTCGTCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTT 102
              | |  |||  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 CCCCCCCCCCCTGGCTCGTCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTT 85

Qy        103 CCTGGGAATGCCTGGGGTTCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGG 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 CCTGGGAATGCCTGGGGTTCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGG 145

Qy        163 ATTATTATTATTATTTTTCTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCC 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 ATTATTATTATTATTTTTCTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCC 205

Qy        223 CTTCTTCTTGCTAAACGCCATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGT 282
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        206 CTTCTTCTTGCTAAATGCCATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGT 265

Qy        283 GGACGTCCGAGGCCACCATCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAG 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GGACGTCCGAGGCCACCATCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAG 325

Qy        343 TCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGC 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 TCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGC 385

Qy        403 CGGCCTCGGTGCCACCCAGCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAG 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 CGGCCTCGGTGCCACCCAGCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAG 445

Qy        463 ACCATGGAGACCCTGCTTGGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGAC 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ACCATGGAGACCCTGCTTGGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGAC 505

Qy        523 GCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 GCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCC 565

Qy        583 TCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTG 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTG 625

Qy        643 GGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTG 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 GGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTG 685

Qy        703 GACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTC 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 GACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTC 745

Qy        763 GAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 GAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACC 805

Qy        823 CTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATC 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 CTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATC 865

Qy        883 GCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAAC 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAAC 925

Qy        943 AACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGC 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 AACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGC 985

Qy       1003 CTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTG 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 CTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTG 1045

Qy       1063 GCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCC 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 GCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCC 1105

Qy       1123 CGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 CGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTT 1165

Qy       1183 GGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGAC 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 GGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGAC 1225

Qy       1243 GATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTG 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 GATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTG 1285

Qy       1303 CGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTG 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 CGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTG 1345

Qy       1363 GTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 GTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTT 1405

Qy       1423 ATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTAT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 ATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTAT 1465

Qy       1483 GACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGC 1542
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1466 GACAATGGCACCCCGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGC 1525

Qy       1543 ATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1526 ATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTG 1585

Qy       1603 CCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 CCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATC 1645

Qy       1663 ACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 ACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAA 1705

Qy       1723 AGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAG 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 AGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAG 1765

Qy       1783 TGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCT 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1766 TGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCT 1825

Qy       1843 GACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAAC 1902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1826 GACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAAC 1885

Qy       1903 AACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCC 1962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1886 AACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCC 1945

Qy      1963 ACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTAC 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1946 ACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTAC 2005

Qy      2023 CCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGG 2082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2006 CCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGG 2065

Qy      2083 GGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTC 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2066 GGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTC 2125

Qy      2143 TGCAACCACGAGGCCCCCAGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACC 2202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2126 TGCAACCACGAGGCCCCCAGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACC 2185

Qy      2203 AACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCG 2262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2186 AACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCG 2245

Qy      2263 CCGAAGGTGGTGGTGCGCAACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGT 2322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2246 CCGAAGGTGGTGGTGCGCAACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGT 2305

Qy      2323 GACTCCTCTTCCTCCAGCTCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCC 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2306 GACTCCTCTTCCTCCAGCTCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCC 2365

Qy      2383 TGGAGGATCCCACCCTCCGCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCC 2442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2366 TGGAGGATCCCACCCTCCGCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCC 2425

Qy      2443 TTCGCCTCCCTGGACCTCAAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCA 2502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2426 TTCGCCTCCCTGGACCTCAAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCA 2485

Qy      2503 GCCGGGAGAGGGGCTGGGACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTG 2562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2486 GCCGGGAGAGGGGCTGGGACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTG 2545

Qy      2563 GGGCCCCCTGCGGCCCGGGCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAA 2622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2546 GGGCCCCCTGCGGCCCGGGCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAA 2605

Qy      2623 CGCAGCCACTCCTTCGACATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCG 2682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2606 CGCAGCCACTCCTTCGACATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCG 2665

Qy      2683 GGCGGCTACAGTCCTCCTCGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTC 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2666 GGCGGCTACAGTCCTCCTCGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTC 2725

Qy      2743 AACGGCATGCTCTTGCCCTTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGC 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2726 AACGGCATGCTCTTGCCCTTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGC 2785

Qy      2803 AGCTCCGAATGGGTGATGGAGAGCACGGTCTAG 2835
             |||||||||||||||||||||||||||||||||
Db      2786 AGCTCCGAATGGGTGATGGAGAGCACGGTCTAG 2818


RESULT 4
AAI58215
ID   AAI58215 standard; cDNA; 2818 BP.
XX
AC   AAI58215;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 418.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,   Liu C,   Asundi V,   Chen R,   Ma Y,   Qian XB,   Ren F,   Wang D;
PI   Wang J,   Wang Z,   Wehrman T,   Xu C,   Xue AJ,   Yang Y,   Zhang J;
PI   Zhao QA,   Zhou P,   Goodrich R,   Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39059.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 418; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 2818 BP; 523 A; 933 C; 825 G; 537 T; 0 other;

  Query Match             88.4%;  Score 2780.2;  DB 22;  Length 2818;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2785;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        43 CACAACCCGGCAGGCTCGTCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTT 102
             | |  |||  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db        26 CCCCCCCCCCCTGGCTCGTCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTT 85

Qy       103 CCTGGGAATGCCTGGGGTTCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGG 162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        86 CCTGGGAATGCCTGGGGTTCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGG 145

Qy       163 ATTATTATTATTATTTTTCTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCC 222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       146 ATTATTATTATTATTTTTCTTTAATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCC 205
```

```
Qy    223 CTTCTTCTTGCTAAACGCCATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGT 282
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    206 CTTCTTCTTGCTAAATGCCATGGATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGT 265

Qy    283 GGACGTCCGAGGCCACCATCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAG 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    266 GGACGTCCGAGGCCACCATCTGCCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAG 325

Qy    343 TCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGC 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    326 TCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGC 385

Qy    403 CGGCCTCGGTGCCACCCAGCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAG 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    386 CGGCCTCGGTGCCACCCAGCCGGGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAG 445

Qy    463 ACCATGGAGACCCTGCTTGGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGAC 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    446 ACCATGGAGACCCTGCTTGGTGGCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGAC 505

Qy    523 GCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCC 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    506 GCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCC 565

Qy    583 TCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTG 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    566 TCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTG 625

Qy    643 GGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTG 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    626 GGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTG 685

Qy    703 GACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTC 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    686 GACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTC 745

Qy    763 GAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACC 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    746 GAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACC 805

Qy    823 CTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATC 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    806 CTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATC 865

Qy    883 GCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAAC 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    866 GCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAAC 925

Qy    943 AACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGC 1002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    926 AACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGC 985

Qy   1003 CTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTG 1062
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    986 CTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTG 1045

Qy   1063 GCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCC 1122
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1046 GCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCC 1105

Qy   1123 CGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTT 1182
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1106 CGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTT 1165

Qy   1183 GGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGAC 1242
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1166 GGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGAC 1225

Qy   1243 GATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTG 1302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1226 GATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTG 1285

Qy   1303 CGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTG 1362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1286 CGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTG 1345

Qy   1363 GTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTT 1422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1346 GTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTT 1405

Qy   1423 ATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTAT 1482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1406 ATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTAT 1465

Qy   1483 GACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGC 1542
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1466 GACAATGGCACCCCGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGC 1525

Qy   1543 ATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTG 1602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1526 ATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTG 1585

Qy   1603 CCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATC 1662
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1586 CCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATC 1645

Qy   1663 ACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAA 1722
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1646 ACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAA 1705

Qy   1723 AGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAG 1782
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1706 AGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAG 1765

Qy   1783 TGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCT 1842
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1766 TGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCT 1825

Qy   1843 GACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAAC 1902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1826 GACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAAC 1885

Qy   1903 AACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCC 1962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1886 AACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCC 1945

Qy   1963 ACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTAC 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1946 ACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTAC 2005

Qy   2023 CCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGG 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2006 CCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGG 2065

Qy   2083 GGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTC 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2066 GGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTC 2125

Qy   2143 TGCAACCACGAGGCCCCCAGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACC 2202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2126 TGCAACCACGAGGCCCCCAGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACC 2185

Qy   2203 AACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2186 AACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCG 2245

Qy   2263 CCGAAGGTGGTGGTGCGCAACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGT 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2246 CCGAAGGTGGTGGTGCGCAACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGT 2305

Qy   2323 GACTCCTCTTCCTCCAGCTCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGACGGGCCCCC 2382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2306 GACTCCTCTTCCTCCAGCTCCCTGGGCAGTGGGGAGGCTGCGGGGCTGGACGGGCCCCC 2365

Qy   2383 TGGAGGATCCCACCCTCCGCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCC 2442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2366 TGGAGGATCCCACCCTCCGCCCCGCGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCC 2425

Qy       2443 TTCGCCTCCCTGGACCTCAAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCA 2502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2426 TTCGCCTCCCTGGACCTCAAGAGTCAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCA 2485

Qy       2503 GCCGGGAGAGGGGCTGGGACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTG 2562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2486 GCCGGGAGAGGGGCTGGGACGTCGGCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTG 2545

Qy       2563 GGGCCCCCTGCGGCCCGGGCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAA 2622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2546 GGGCCCCCTGCGGCCCGGGCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAA 2605

Qy       2623 CGCAGCCACTCCTTCGACATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCG 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2606 CGCAGCCACTCCTTCGACATGGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCG 2665

Qy       2683 GGCGGCTACAGTCCTCCTCGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTC 2742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2666 GGCGGCTACAGTCCTCCTCGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTC 2725

Qy       2743 AACGGCATGCTCTTGCCCTTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGC 2802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2726 AACGGCATGCTCTTGCCCTTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGC 2785

Qy       2803 AGCTCCGAATGGGTGATGGAGAGCACGGTCTAG 2835
              |||||||||||||||||||||||||||||||||
Db       2786 AGCTCCGAATGGGTGATGGAGAGCACGGTCTAG 2818


RESULT 5
AAS77723
ID   AAS77723 standard; cDNA; 969 BP.
XX
AC   AAS77723;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #13527.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG13536.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 13527; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
```

```
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 969 BP; 166 A; 330 C; 316 G; 157 T; 0 other;

  Query Match              30.8%;  Score 969;  DB 23;  Length 969;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-194;
  Matches  969;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1867 ATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTAC 1926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTAC 60

Qy       1927 GACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATC 1986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATC 120

Qy       1987 GTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGC 2046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGC 180

Qy       2047 CAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTG 2106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTG 240

Qy       2107 GTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAGCAAG 2166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAGCAAG 300

Qy       2167 ATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCA 2226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCA 360

Qy       2227 AGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGCAACGAG 2286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGCAACGAG 420

Qy       2287 CTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGCTCCCTG 2346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGCTCCCTG 480

Qy       2347 GGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACCCTCCGCCCCG 2406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACCCTCCGCCCCG 540

Qy       2407 CGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTCAAGAGT 2466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTCAAGAGT 600

Qy       2467 CAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGGACGTCG 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGGACGTCG 660

Qy       2527 GCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGGGCCAGG 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGGGCCAGG 720

Qy       2587 AGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGACATGGGG 2646
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 AGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGACATGGGG 780

Qy        2647 GACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCTCGGAAG 2706
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCTCGGAAG 840

Qy        2707 GTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCCTTTGAG 2766
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 GTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCCTTTGAG 900

Qy        2767 GAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATGGAGAGC 2826
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 GAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATGGAGAGC 960

Qy        2827 ACGGTCTAG 2835
               |||||||||
Db         961 ACGGTCTAG 969


RESULT 6
AAI60001
ID   AAI60001 standard; cDNA; 771 BP.
XX
AC   AAI60001;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 3990.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM40845.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 3990; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 771 BP; 144 A; 259 C; 197 G; 171 T; 0 other;

  Query Match             21.6%;  Score 678.8;  DB 22;  Length 771;
  Best Local Similarity   95.8%;  Pred. No. 1.8e-133;
  Matches  730;  Conservative    0;  Mismatches   27;  Indels    5;  Gaps    3;

Qy         529 CCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAG 588
               ||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           8 CCCCAAGACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAG 67

Qy         589 GGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGC 648
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          68 GGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGC 127

Qy         649 AACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTG 708
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         128 AACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTG 187

Qy         709 ACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGC 768
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         188 ACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGC 247

Qy         769 CTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGG 828
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         248 CTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGG 307

Qy         829 GGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGAT 888
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         308 GGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGAT 367

Qy         889 GAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTC 948
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         368 GAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTC 427

Qy         949 CAT-GGCCTGCCGTGGGACTCCGT-GCGACGCATGG---TCAACCTCCACCAGCTGAGCC 1003
               ||| |||||||||||||||||||| |||||||||||     | | |||||||||||||||
Db         428 CATGGGCCTGCCGTGGGACTCCGTGGCGACGCATGGGTTCAACCTTCCACCAGCTGAGCC 487

Qy        1004 TGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGG 1063
               |||||||||||||| |||||||||||||    ||||||||||||||||||||||||||||
Db         488 TGGACCACAACCTGGTGGATCACATCGCTCGAGGCACCTTTGCAGACCTGCAGAAACTGG 547

Qy        1064 CCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCC 1123
               |||||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||
Db         548 CCCGCCTGGATCTCACCTCCAATCGGTTGCATAAGCTGCCCCCTGATCCCATCTTTGCCC 607

Qy        1124 GCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTG 1183
               ||||||||||||| |||||||||| ||||||||||||||||||||||||||||| |||||
Db         608 GCTCCCAGGCTTCCGCTTTGACAGACACACCCTTTGCCCCACCCTTGTCCTTTATTTTTG 667

Qy        1184 GGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACG 1243
               ||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||||
Db         668 GGGGTAACCCACTTCACTGCAATTGAGAGCTTCTCTGGCTGCGGAGGCTCGATCGGGACG 727

Qy        1244 ATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTC 1285
               |   ||||||||| ||||| || |||| |||| |||||||||
Db         728 AAACCCTGGAAACTTGTGGTTCTCCAGTGGGCGTCAAGGGTC 769


RESULT 7
AAH25568
```

```
ID   AAH25568 standard; DNA; 1887 BP.
XX
AC   AAH25568;
XX
DT   05-SEP-2001  (first entry)
XX
DE   Nucleotide sequence of an interferon omega-1 like protein NOV2.
XX
KW   Interferon omega-1 like protein; Nov2; membrane bound protein;
KW   secreted protein;. spermatogenesis; male infertility; neoplasia;
KW   blood circulation; immunological disorder; autoimmune disease;
KW   inflammatory disease; cardiovascular disease; metabolic disease;
KW   cancer; viral infection; acute lymphoblastic leukemia; glioma;
KW   neurological disease; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disorder; hematopoietic disorder; neurodegenerative disorder;
KW   immune disorder; hematopoietic disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1887
FT                   /*tag=  a
FT                   /product= "interferon omega-1 like protein NOV2"
XX
PN   WO200142471-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US33463.
XX
PR   09-DEC-1999;   99US-0169887.
PR   10-DEC-1999;   99US-0170230.
PR   07-DEC-2000; 2000US-0170230.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Prayaga SK,  Shimkets RA;
XX
DR   WPI; 2001-381691/40.
DR   P-PSDB; AAB84469.
XX
PT   Novel polypeptides designated as NOV polypeptides, useful in detection,
PT   prevention and treatment of a broad range of pathological states  -
XX
PS   Claim 8; Page 15-16; 92pp; English.
XX
CC   The present sequence encodes an interferon omega-1 like protein,
CC   designated NOV2. The protein is a membrane bound and secreted protein.
CC   NOV polypeptides and polynucleotides are useful for treating or
CC   preventing a NOV-associated disorder, and in the manufacture of a
CC   medicament for treating a syndrome associated with a human disease such
CC   as NOV-associated disorder. They are also useful for treating disorders
CC   or syndromes including those involved in development, differentiation
CC   and activation of thymic immune cells, pathologies related to
CC   spermatogenesis and male infertility, diagnosis of several human
CC   neoplasias, in disease or pathologies of cells in blood circulation
CC   such as red blood cells and platelets, various immunological disorders
CC   and/or pathologies, autoimmune and inflammatory diseases,
CC   cardiovascular diseases, metabolic diseases, cancer growth and
CC   metastasis, viral infections, cancer therapy, acute lymphoblastic
CC   leukemia, in gliomas, neurological diseases, neurodegenerative disorders,
CC   Alzheimer's disease, Parkinson's disorder, and hematopoietic disorders.
CC   NOV polypeptides and polynucleotides are also useful for treating
CC   neurodegenerative disorders, immune disorders and hematopoietic
CC   disorders.
XX
SQ   Sequence 1887 BP; 259 A; 736 C; 577 G; 315 T; 0 other;

  Query Match             20.2%;  Score 634.6;  DB 22;  Length 1887;
  Best Local Similarity   61.3%;  Pred. No. 4.6e-124;
  Matches 1079;  Conservative    0;  Mismatches  664;  Indels   16;  Gaps    3;

Qy         507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
               ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db          63 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 122

Qy         567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                 |    ||||||||       || || || || || ||      | |||| ||| |
Db         123 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 182

Qy         627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
               ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db         183 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 242

Qy         687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
               |||||| || ||| || | |||| ||||||  |||||||||||| ||||| |     ||
Db         243 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 302

Qy         747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                 ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db         303 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 362

Qy         807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
                || || |||| |   || || ||||||||||||| |||  ||||| ||| | | ||||||
Db         363 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 422

Qy         867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
               ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db         423 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 482

Qy         927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
                ||||||||||||||||||| |    ||||| ||||| || || | ||| ||| |||
Db         483 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 542

Qy         987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
               | || ||  | || |||| |||||||||||||||| |    |  |||  ||| | ||| |
Db         543 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 602

Qy        1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                  |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db         603 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 662

Qy        1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
                || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db         663 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 722

Qy        1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
               | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db         723 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 782

Qy        1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                 | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db         783 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 842

Qy        1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
               ||||||||||   ||| | |||||||||||||| ||||||||||||  |  | || || |
Db         843 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 902

Qy        1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
               | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db         903 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 962

Qy        1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
               ||||||    |||| | | |  |||||  | ||| | | ||| ||| ||||| |||||
Db         963 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1022

Qy        1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                 |  ||   | ||    |||||| || ||||||  |  |  |||||    |    ||  |
Db        1023 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1082

Qy        1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
               |||   |||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db        1083 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1142

Qy        1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                 |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
```

```
Db        1143 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1202

Qy        1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                  ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db        1203 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1262

Qy        1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
                |      | | ||| |        | || || |  || |  ||| ||||        ||
Db        1263 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1314

Qy        1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
               ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db        1315 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1374

Qy        1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
                | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db        1375 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1434

Qy        1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                   | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db        1435 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1494

Qy        1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
                |||  || ||||  ||||||   |||||||||||       |||||||||||||  |||
Db        1495 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1554

Qy        2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
               | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db        1555 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 1614

Qy        2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
               |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db        1615 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 1674

Qy        2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
                |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db        1675 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 1734

Qy        2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
               || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db        1735 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 1794

Qy        2239 GCCGGGGCCCCGCCGCAGG 2257
                |||   |   |||   ||
Db        1795 CCCGCCCCGGAGCCCGCGG 1813

RESULT 8
AAH78215
ID   AAH78215 standard; DNA; 1887 BP.
XX
AC   AAH78215;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Nucleotide sequence of a human secreted polypeptide.
XX
KW   Human; secreted polypeptide; nervous disease; muscular disease; tumour;
KW   gastrointestinal ulceration; spinal cord disease; trachea disease;
KW   thyroid gland disease; ovary disease; prostate disease; heart disease;
KW   renal gland disease; small intestine disease; thymus disease;
KW   lymph node disease; muscular system disease; colon disease;
KW   lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
KW   myocardial infarction; angioplasty; liver disease; coagulation disorder;
KW   microbial disease; immune disorder; inflammation; transplant rejection;
KW   bone thickness; bone density; ferroxidase loss; apoptosis;
KW   vascular smooth cell proliferation; vaccine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1887
FT                   /*tag=  a
FT                   /product= "secreted polypeptide"
XX
PN   WO200166690-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US07143.
XX
PR   06-MAR-2000; 2000US-0187107.
PR   13-MAR-2000; 2000US-0188916.
PR   03-OCT-2000; 2000US-0236874.
PR   03-OCT-2000; 2000US-0237846.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Agarwal P,  Murdoch PR,  Rizvi SK,  Smith RF,  Xiang Z;
XX
DR   WPI; 2001-570768/64.
DR   P-PSDB; AAG67523.
XX
PT   Novel isolated secreted polypeptide useful for treating nervous and
PT   muscular diseases, gastrointestinal ulceration, coagulation and immune
PT   disorders, microbial diseases, inflammation and transplant rejection -
XX
PS   Claim 2; Page 49-50; 102pp; English.
XX
CC   The present sequence encodes a human secreted polypeptide. The
CC   secreted polypeptides and polynucleotides are useful for treating
CC   nervous and muscular diseases, for inhibiting tumour formation and
CC   metastasis, for treating gastrointestinal ulceration, for preventing
CC   and treating diseases in spinal cord, thyroid gland, ovary, prostate,
CC   renal gland, small intestine, heart, trachea, thymus, lymph node,
CC   muscular system and colon, for treating lipase deficiency in cystic
CC   fibrosis and pancreatitis, for treating undesirable clot formation
CC   such as myocardial infarction, during angioplasty and all surgical
CC   procedures that require decreased blood clot formation, for treating
CC   liver diseases, coagulation disorders and microbial diseases, for
CC   treating immune disorders, for treating inflammation and transplant
CC   rejection, for enhancing bone thickness and increasing bone density,
CC   for reducing the loss of essential ferroxidases, for suppressing
CC   apoptosis, and for regulating vascular smooth cell proliferation. They
CC   may also be used as vaccines.
XX
SQ   Sequence 1887 BP; 259 A; 736 C; 577 G; 315 T; 0 other;

  Query Match             20.2%;  Score 634.6;  DB 22;  Length 1887;
  Best Local Similarity   61.3%;  Pred. No. 4.6e-124;
  Matches 1079;  Conservative    0;  Mismatches  664;  Indels   16;  Gaps    3;

Qy         507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
               ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db          63 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 122

Qy         567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                 |    ||||||||       || || || || || || ||     | ||||| ||| |
Db         123 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 182

Qy         627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
               ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db         183 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 242

Qy         687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
               |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db         243 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 302

Qy         747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                 ||||    |||||   |  ||| ||  ||||||| || ||  |||| ||||||  |
Db         303 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 362

Qy         807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
                || || ||||  |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db         363 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 422
```

```
Qy       867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
             ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db       423 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 482

Qy       927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
              ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db       483 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 542

Qy       987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
             | || ||  | || |||| |||||||||||||||| | |  |  |||  ||| | ||| |
Db       543 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 602

Qy      1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db       603 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 662

Qy      1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
              || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db       663 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 722

Qy      1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
             | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db       723 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 782

Qy      1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
               | || | |||||| || |||| ||  |||| |  ||||||   |  ||    || ||
Db       783 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 842

Qy      1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
             ||||||||||   ||| | ||||||||||||||| |||||||||||  |  | || || | 
Db       843 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 902

Qy      1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
             | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db       903 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 962

Qy      1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
             ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||   
Db       963 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1022

Qy      1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
               |  ||   | ||    ||||||| || ||||||  |  |  |||||    |    ||  |
Db      1023 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1082

Qy      1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
             |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db      1083 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1142

Qy      1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
               |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db      1143 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1202

Qy      1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                ||    ||      | ||  | ||| |     | |||   |     |     |||| 
Db      1203 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1262

Qy      1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
              |      | | ||| |        | || || |  || |  ||| ||||        ||
Db      1263 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1314

Qy      1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
             ||  | || || ||| |||||| | |       | |    ||  |  |     |||||||||
Db      1315 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1374

Qy      1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
              | |||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db      1375 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1434

Qy      1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                 | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db      1435 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1494

Qy      1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
              |||  || ||||  ||||||   |||||||||||       ||||||||||||||  |||
Db      1495 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1554

Qy      2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
             | |||| |  |||     | ||  |||  | |   ||||   |   | ||||||||||| 
Db      1555 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 1614

Qy      2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
             |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | || 
Db      1615 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 1674

Qy      2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
              |  |||||||||||||   |  |  | || ||||| |        ||||||  | ||| | 
Db      1675 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 1734

Qy      2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
             || ||||  || | ||| |||||||||||||||||| |   ||  | || |    ||| || 
Db      1735 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 1794

Qy      2239 GCCGGGGCCCCGCCGCAGG 2257
              |||   |   |||   ||
Db      1795 CCCGCCCCGGAGCCCGCGG 1813


RESULT 9
AAI66986
ID   AAI66986 standard; cDNA; 1887 BP.
XX
AC   AAI66986;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Human LRR family member, 33395 coding sequence.
XX
KW   LRR; leucine rich repeat; 33395; cytostatic; anti-HIV; antidiabetic;
KW   antiarthritic; neuroprotective; dermatological; immunosuppresive; ss;
KW   antiinflammatory; antiasthmatic; antiulcer; antianaemic; hepatotropic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1887
FT                   /*tag=  a
FT                   /product= "33395"
XX
PN   WO200172827-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US09470.
XX
PR   24-MAR-2000; 2000US-191863P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Glucksmann MA;
XX
DR   WPI; 2001-626254/72.
DR   P-PSDB; AAG65805.
XX
PT   New polypeptide for preventing or treating disorders associated with
PT   cellular adhesion, proliferation or differentiation, comprises
PT   polypeptide 33395, a member of the leucine rich repeat protein family
PT   -
XX
PS   Claim 1; Fig 1; 133pp; English.
XX
CC   The invention provides an isolated nucleic acid encoding a polypeptide
CC   of the leucine rich repeat (LRR) family, designated 33395. The 33395
CC   polypeptide can be expressed by standard recombinant methodology. The
CC   33395 polynucleotides and polypeptide can be used to prevent or treat
CC   disorders associated with 33395 expression, for example those involving
```

```
CC   aberrant cellular adhesion, proliferation or differentiation. Specific.
CC   examples include meloblastoma, juvenile AIDS, diabetes mellitus,
CC   rheumatoid arthritis, systemic lupus erythematosis, multiple sclerosis,
CC   Chron's disease, ulcerative colitis, asthma, anemia, and  chronic active
CC   hepatitis. The present sequence represents the coding sequence of the
CC   human 33395 polypeptide.
XX
SQ   Sequence 1887 BP; 257 A; 735 C; 579 G; 316 T; 0 other;

  Query Match            20.1%;  Score 633;  DB 22;  Length 1887;
  Best Local Similarity  61.3%;  Pred. No. 1e-123;
  Matches 1078;  Conservative    0;  Mismatches  665;  Indels   16;  Gaps    3;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566

Db         63 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 122

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626

Db        123 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 182

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686

Db        183 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 242

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746

Db        243 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 302

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806

Db        303 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 362

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866

Db        363 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 422

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926

Db        423 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 482

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986

Db        483 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 542

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046

Db        543 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 602

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106

Db        603 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 662

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166

Db        663 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 722

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226

Db        723 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 782

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286

Db        783 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 842

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346

Db        843 CTACTTTTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 902

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406

Db        903 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 962

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466

Db        963 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1022

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526

Db       1023 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1082

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586

Db       1083 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1142

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644

Db       1143 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1202

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704

Db       1203 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1262

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764

Db       1263 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1314

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824

Db       1315 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1374

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884

Db       1375 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1434

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944

Db       1435 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1494

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004

Db       1495 GTGTATGAGGGCAGTGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1554

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064

Db       1555 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 1614

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124

Db       1615 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 1674

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178

Db       1675 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCGGCAAGGCCAAGATTCCCGCGCCC 1734

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238

Db       1735 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 1794

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257

Db       1795 CCCGCCCCGGAGCCCGCGG 1813


RESULT 10
AAI66985
ID   AAI66985 standard; cDNA; 2637 BP.
XX
AC   AAI66985;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Human leucine-rich repeat (LRR) family member, 33395 cDNA sequence.
XX
KW   LRR; leucine rich repeat; 33395; cytostatic; anti-HIV; antidiabetic;
KW   antiarthritic; neuroprotective; dermatological; immunosuppresive; ss;
```

```
KW   antiinflammatory; antiasthmatic; antiulcer; antianaemic; hepatotropic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..434
FT                   /*tag=  a
FT   CDS             435..2321
FT                   /*tag=  b
FT                   /product= "33395"
FT                   /note= "the coding sequence is specifically claimed"
FT   3'UTR           2322..2637
FT                   /*tag=  c
XX
PN   WO200172827-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US09470.
XX
PR   24-MAR-2000; 2000US-191863P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Glucksmann MA;
XX
DR   WPI; 2001-626254/72.
DR   P-PSDB; AAG65805.
XX
PT   New polypeptide for preventing or treating disorders associated with
PT   cellular adhesion, proliferation or differentiation, comprises
PT   polypeptide 33395, a member of the leucine rich repeat protein family
PT   -
XX
PS   Claim 1; Fig 1; 133pp; English.
XX
CC   The invention provides an isolated nucleic acid encoding a polypeptide
CC   of the leucine rich repeat (LRR) family, designated 33395. The 33395
CC   polypeptide can be expressed by standard recombinant methodology. The
CC   33395 polynucleotides and polypeptide can be used to prevent or treat
CC   disorders associated with 33395 expression, for example those involving
CC   aberrant cellular adhesion, proliferation or differentiation. Specific
CC   examples include meloblastoma, juvenile AIDS, diabetes mellitus,
CC   rheumatoid arthritis, systemic lupus erythematosis, multiple sclerosis,
CC   Chron's disease, ulcerative colitis, asthma, anemia, and  chronic active
CC   hepatitis. The present sequence represents the cDNA encoding the human
CC   33395 polypeptide.
XX
SQ   Sequence 2637 BP; 452 A; 976 C; 758 G; 451 T; 0 other;

  Query Match             20.1%;  Score 633;  DB 22;  Length 2637;
  Best Local Similarity   61.3%;  Pred. No. 1.1e-123;
  Matches 1078;  Conservative    0;  Mismatches  665;  Indels   16;  Gaps    3;

Qy         507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
               ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db         497 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 556

Qy         567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                 |    ||||||||       || || || || || || ||     | ||||| ||| |
Db         557 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 616

Qy         627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
               ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db         617 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 676

Qy         687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
               |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db         677 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 736

Qy         747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                 ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db         737 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 796

Qy         807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
                || || |||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db         797 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 856

Qy         867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
               ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db         857 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 916

Qy         927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
                ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db         917 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 976

Qy         987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
               | || ||  | || |||| |||||||||||||||| |    |  |||  ||| | ||| |
Db         977 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 1036

Qy        1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                  |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db        1037 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1096

Qy        1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
                || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db        1097 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1156

Qy        1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
               | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db        1157 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1216

Qy        1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                 | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db        1217 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1276

Qy        1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
               |||||| |||   ||| | |||||||||||||| ||||||||||||  |  | || || |
Db        1277 CTACTTTTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1336

Qy        1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
               | |||       ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db        1337 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1396

Qy        1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
               ||||||    |||| | | |  |||||  | ||| | | ||| ||| ||||| |||||
Db        1397 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1456

Qy        1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                 |  ||   | ||    |||||| || |||||  |  |   |||||    |    ||  |
Db        1457 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1516

Qy        1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
               |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db        1517 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1576

Qy        1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                 |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db        1577 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1636

Qy        1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                  ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db        1637 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1696

Qy        1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
                |      | | ||| |        | || || |  || |  ||| ||||        ||
Db        1697 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1748

Qy        1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
               ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db        1749 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1808

Qy        1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
                | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db        1809 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1868
```

```
Qy   1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
              | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db   1869 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1928

Qy   1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
           |||  || | ||  |||||||   |||||||||||       |||||||||||||  |||
Db   1929 GTGTATGAGGGCAGTGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1988

Qy   2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
          | |||| |  |||     | ||  |||  | |   ||||   |   | ||||||||||| 
Db   1989 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2048

Qy   2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
          |||||| | |   | || ||| ||||||| ||| || | ||||||||||| || | || 
Db   2049 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2108

Qy   2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
           |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| ||
Db   2109 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCC 2168

Qy   2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
          || ||||  || | ||| ||||||||||||||||| |   ||  | || |    ||| || 
Db   2169 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2228

Qy   2239 GCCGGGGCCCCGCCGCAGG 2257
           |||   |   |||   ||
Db   2229 CCCGCCCCGGAGCCCGCGG 2247

RESULT 11
AAD38695
ID   AAD38695 standard; cDNA; 2369 BP.
XX
AC   AAD38695;
XX
DT   23-SEP-2002  (first entry)
XX
DE   Human LP220 secreted protein encoding cDNA.
XX
KW   Human; secreted protein; atherosclerosis; Alzheimer's disease; LP220;
KW   diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
KW   rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
KW   reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
KW   gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;
KW   chomosome 11q13; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             167..2074
FT                   /*tag=  a
FT                   /product= "Human LP220 secreted protein"
FT   sig_peptide     167..214
FT                   /*tag=  b
FT   mat_peptide     215..2071
FT                   /*tag=  c
FT                   /product= "Mature human LP220 secreted protein"
XX
PN   WO200226801-A2.
XX
PD   04-APR-2002.
XX
PF   14-SEP-2001; 2001WO-US26026.
XX
PR   28-SEP-2000; 2000US-236088P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Su EW,  Wang H;
XX
DR   WPI; 2002-471259/50.
DR   P-PSDB; AAE23980.
XX
```

```
PT   Novel proteins and polynucleotides of secreted proteins useful for
PT   treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,
PT   diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury
PT   -
XX
PS   Claim 1; Page 124-127; 145pp; English.
XX
CC   The invention relates to human secreted polypeptides designated LP095,
CC   LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic
CC   acid molecules encoding such polypeptides. Novel secreted proteins of
CC   the invention are used for treating diseases such as atherosclerosis,
CC   Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,
CC   arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe
CC   combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,
CC   reperfusion injury, neoplasms and cancer especially liver cancer. They
CC   are also used for wound healing. Polynucleotides of the invention can
CC   be used to generate transgenic animals or knock out animals, which in
CC   turn, are useful in the development and screening of therapeutically
CC   useful reagents for use in the treatment of diseases associated with
CC   LP polypeptide associated activity. They are also used in gene therapy.
CC   The present sequence is human LP220 secreted protein encoding cDNA.
CC   LP220 gene is located on chomosome 11q13.
XX
SQ   Sequence 2369 BP; 342 A; 855 C; 755 G; 417 T; 0 other;

  Query Match              19.8%;  Score 622.6;  DB 24;  Length 2369;
  Best Local Similarity    62.3%;  Pred. No. 1.6e-121;
  Matches 1089;  Conservative    0;  Mismatches  609;  Indels   51;  Gaps    5;

Qy    520 GACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGC 579
          | |||||||||   |  ||| ||||||||||| ||||| ||||| ||  | ||||| || 
Db    209 GCCGCCTGCCCGCTGCCCTGCGTCTGCCAGAACCTGTCCGAGTCGCTCAGCACCCTCTGT 268

Qy    580 CCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGC 639
           ||  |   || ||||| |||||| || ||  |  | ||||||||| ||||||||||||| 
Db    269 GCCCACCGAGGCCTGCTGTTTGTGCCGCCCAACGTGGACCGGCGCACAGTGGAGCTGCGG 328

Qy    640 CTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTG 699
          ||||  | ||||||||||     | |  |  | ||  |||||   ||||||||||||| |||
Db    329 CTGGCTGACAACTTCATCCAGGCCCTGGGGCCCCCTGACTTCCGCAACATGACGGGACTG 388

Qy    700 GTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGAC 759
          ||||||||||| |||||  | ||  |||||| || |||   | ||    |||||  ||||
Db    389 GTGGACCTGACACTGTCTCGCAATGCCATCACCCGCATTGGGGCCCGCGCCTTTGGGGAC 448

Qy    760 CTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGAC 819
          ||||||||||| || ||||| || |||||| ||||  |||||      || ||    |  
Db    449 CTCGAGAGCCTGCGTTCCCTCCACCTTGACGGCAACAGGCTGGTGGAGCTGGGCACCGGG 508

Qy    820 ACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGC 879
          | |||||||||||  ||||| ||||||||||| ||| | | |  ||||||||||||| ||
Db    509 AGCCTCCGGGGCCCCGTCAATCTGCAGCACCTCATCCTCAGCGGCAACCAGCTGGGCCGC 568

Qy    880 ATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTAC 939
          |||||    |  || || || ||||||||   ||   |||||||| |||||||| ||||||
Db    569 ATCGCGCCGGGAGCCTTCGACGACTTCCTAGAGAGCCTGGAGGACCTGGACCTGTCCTAC 628

Qy    940 AACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTG 999
          ||||||||||      |||| |||| |  | |  |   ||||     ||| |||   || 
Db    629 AACAACCTCCGGCAGGTGCCCTGGGCCGGCATCGGCGCCATGCCTGCCCTGCACACCCTC 688

Qy   1000 AGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAA 1059
          | ||||||||| |||||  | ||     |  ||   |||  |||| ||  | ||     | 
Db    689 AACCTGGACCATAACCTTATTGACGCACTGCCCCCAGGCGCCTTCGCCCAGCTCGGTCAG 748

Qy   1060 CTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTT 1119
          ||  |||||||||| ||||||||||| || |||   | |||| | || ||  ||  | || 
Db    749 CTCTCCCGCCTGGACCTCACCTCCAACCGCCTGGCCACGCTGGCTCCGGACCCGCTTTTC 808

Qy   1120 GCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGT 1179
           | ||    |  | |            ||     ||| | || | ||| ||    | || 
Db    809 TCTCGTGGGCGTGAT------------GCAGAGGCCTCTCCCGCCCCCCTGGTGCTGAGC 856
```

```
Qy     1180 TTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGG 1239
            ||| | || ||||| || |||||||| |||||||| || ||||||||| |||| | ||||
Db      857 TTTAGCGGGAACCCCCTGCACTGCAACTGTGAGCTGCTGTGGCTGCGGCGGCTGGCGCGG 916

Qy     1240 GACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCAT 1299
               || ||||||||||| || | ||||||    |||||    || ||||||||||||
Db      917 CCGGACGACCTGGAAACGTGCGCCTCCCCGCCCGGCCTGGCCGGCCGCTACTTCTGGGCA 976

Qy     1300 GTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTG 1359
            ||||  ||||  |||||    || |||||||| |||||  |||  ||||| ||     |
Db      977 GTGCCCGAGGGCGAGTTCTCCTGTGAGCCGCCCCTCATTGCCCGCCACACGCAGCGCCTC 1036

Qy     1360 CTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCC 1419
              ||| ||||| ||||||  |||||| ||   ||||   ||| | || ||||||   ||
Db     1037 TGGGTGCTGGAAGGCCAGCGGGCCACGCTGCGGTGCCGGGCCCTGGGTGACCCCGCGCCT 1096

Qy     1420 CTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTC 1479
               || |||||||| |  || || |||||  |||| || ||||||||  |  ||   |
Db     1097 ACCATGCACTGGGTCGGTCCTGACGACCGGTTGGTTGGCAACTCCTCCCGAGCCCGGGCT 1156

Qy     1480 TATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 1539
            |    ||| || ||| | || ||     | |||    ||  ||||  ||| | || ||||
Db     1157 TTCCCCAACGGGACCTTAGAGATTGGGGTGACCGGCGCTGGGGACGCTGGGGGCTACACC 1216

Qy     1540 TGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAG 1599
            ||||| ||  ||||  |||| || |||||||| |||   || ||  |     |  |
Db     1217 TGCATCGCCACCAACCCTGCTGGTGAGGCCACAGCCCGAGTAGAACTGCGGGTGCTGGCC 1276

Qy     1600 CTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGAC 1659
             |||| ||     | |||||||    |      |  | ||      || | | ||| |||
Db     1277 TTGCCCCATGGTGGGAACAGCAGTGCCGAGGGGGGCCGCC------CCGGGCCCTCGGAC 1330

Qy     1660 ATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCC 1719
            ||| | | ||||      ||       |  ||||  || |    ||     |||||
Db     1331 ATCGCCGCCTCCGCTCGCACTGCTGCCGAGGGTGAGGGGACGCTGGAGTCTGAGCC---- 1386

Qy     1720 AAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTC 1779
                              || ||||  ||| | || |||||| ||||||| |  |||||
Db     1387 -----------------AGCCGTGCAGGTGACGGAGGTGACCGCCACCTCAGGGCTGGTG 1429

Qy     1780 AAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGC 1839
            |  |||  |  | |   | ||||   ||  |||  ||||| |||  | ||||||||| ||
Db     1430 AGCTGGGGTCCCGGGCGGCCAGCCGACCCAGTGTGGATGTTCCAAATCCAGTACAACAGC 1489

Qy     1840 TCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTC 1899
               || ||||||   || || ||| ||||  |||||||||||| | |   |||| || |
Db     1490 AGCGAAGATGAGACCCTCATCTACCGGATTGTCCCAGCCTCCAGCCACCACTTCCTGCTG 1549

Qy     1900 AACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACA 1959
            ||  |||| ||  | ||  ||| ||| ||| | ||  |||||||| |||     | | |
Db     1550 AAGCACCTCGTCCCCGGCGCTGACTATGACCTCTGCCTGCTGGCCTTGTCACCGGCCGCT 1609

Qy     1960 G---CCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCT 2016
            |   || |    |||||||||||||   |  ||||||| ||||| |||| |||   | |
Db     1610 GGGCCCTCTGACCTCACGGCCACCAGGCTGCTGGGCTGTGCCCATTTCTCCACGCTGCCG 1669

Qy     2017 GACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTC 2076
            | ||  || | ||||||  || ||||   |||  | |||||||| ||| ||| | ||| |
Db     1670 GCCTCGCCCCTGTGCCACGCCCTGCAGGCCCACGTGCTGGGCGGGACCCTGACCGTGGCC 1729

Qy     2077 ATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTCATGGTGCGCTAC 2136
             | |||||  |  | |||||  |  | ||||||||||| ||   | |  |||| ||   |
Db     1730 GTGGGGGGTGTGCTGGTGGCTGCCTTACTGGTCTTCATTGTGGCCTTGCTGGTTCGGGGC 1789

Qy     2137 AAGGTCTGCAACCACGAGGCCCCCAGCAAGATGGCAGCGGCCGTGAGCAATGTGTACTCG 2196
              ||    |    | |    || |  |   | |          | ||| | ||  | ||
Db     1790 CGGGGGGCCGGAAATGGCCGCCTCCCCCTCAAG---------CTCAGCCACGTCCAGTCC 1840

Qy     2197 CAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAG 2256
            |||||||| || | ||     ||  | || |    |   ||  |  ||  ||| ||||
Db     1841 CAGACCAATGGAGGCCCCAGCCCCACACCCAAGGCCCACCCGCCGCGGAGCCCCCCGCCC 1900

Qy     2257 GGCCCGCCG 2265
             | |||| |
Db     1901 CGGCCGCAG 1909
```

```
RESULT 12
ABK70010
ID   ABK70010 standard; DNA; 2855 BP.
XX
AC   ABK70010;
XX
DT   15-JUL-2002  (first entry)
XX
DE   cDNA encoding human Pro peptide #50.
XX
KW   Human; ss; gene; PRO; secreted protein; transmembrane protein;
KW   genetic disorder; tumour; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200224888-A2.
XX
PD   28-MAR-2002.
XX
PF   29-AUG-2001; 2001WO-US27099.
XX
PR   01-SEP-2000; 2000US-229896P.
PR   05-SEP-2000; 2000US-230621P.
PR   22-SEP-2000; 2000US-235147P.
PR   10-NOV-2000; 2000WO-US30873.
PR   12-JAN-2001; 2001US-261878P.
PR   16-JAN-2001; 2001US-261910P.
PR   16-JAN-2001; 2001US-261939P.
PR   16-JAN-2001; 2001US-262150P.
PR   25-JAN-2001; 2001US-264395P.
PR   02-FEB-2001; 2001US-266421P.
PR   09-FEB-2001; 2001US-267623P.
PR   28-FEB-2001; 2001WO-US06520.
PR   09-MAR-2001; 2001US-274399P.
PR   03-APR-2001; 2001US-280982P.
PR   04-APR-2001; 2001US-282129P.
PR   04-APR-2001; 2001US-282199P.
PR   09-MAY-2001; 2001US-290589P.
PR   25-MAY-2001; 2001WO-US17092.
PR   01-JUN-2001; 2001WO-US17800.
PR   20-JUN-2001; 2001WO-US19692.
PR   29-JUN-2001; 2001WO-US21066.
PR   09-JUL-2001; 2001WO-US21735.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Eaton DL,  Filvaroff E,  Goddard A,  Grimaldi JC;
PI   Gurney AL,  Smith V,  Stephan J,  Watanabe CK,  Wood WI,  Zhang Z;
PI   Fong S;
XX
DR   WPI; 2002-362426/39.
DR   P-PSDB; ABG34079.
XX
PT   New PRO polypeptides and polynucleotides encoding the polypeptides,
PT   useful in gene therapy, chromosome identification, tissue typing, or
PT   for genetic analysis of individuals with genetic disorders  -
XX
PS   Claim 2; Figure 99; 218pp; English.
XX
CC   This invention relates to the cDNA and protein sequences of novel
CC   secreted and transmembrane polypeptides PRO polypeptides. The
CC   invention also comprises a method for producing the proteins of the
CC   invention by recombinant means and antibodies specific for the protein
CC   of the invention. The antibody may be used for detecting the PRO
CC   proteins of the invention and may be used to modify their activity.
CC   polynucleotides may be used as hybridisation probes for a cDNA library
CC   to isolate the full-length PRO cDNA or to isolate other cDNAs, to
CC   construct hybridisation probes for mapping the gene which encodes that
CC   PRO and for genetic analysis of individuals with genetic disorders, in
```

```
CC   assays to identify other proteins or molecules involved in binding
CC   reaction, to generate transgenic animals or knock-out animals which in
CC   turn are useful in the development and screening of therapeutically
CC   useful reagents, for chromosome identification, and tissue typing. The
CC   PRO polypeptides are useful in gene therapy, and as molecular weight
CC   markers for protein electrophoresis purposes. The sequences may
CC   also be used to detect overexpression on PRO polypeptides in cancerous
CC   tumours and for screening for differentially expressed genes using
CC   microarray technology. The present sequence represents a cDNA encoding
CC   a human PRO protein of the invention.
XX
SQ   Sequence 2855 BP; 552 A; 1007 C; 799 G; 497 T; 0 other;

  Query Match             19.8%;  Score 621.8;  DB 24;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 2.5e-121;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy         507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
               ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db         798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy         567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                 |    ||||||||       || || || || || ||     | |||| ||| |
Db         858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy         627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
               ||  |||||||| |||   | |||||||||||  |  | |  ||||||  ||| | |||||
Db         918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy         687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
               |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db         978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy         747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                 ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db        1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy         807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
                || || ||| | |   || || ||||||||||||| |||  ||||| ||| | | ||||||
Db        1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy         867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
               ||||||||  |   | || |  |  ||  |  || || || |       |||| ||||||| ||
Db        1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy         927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
                ||||||||||||||||||||| |    ||||| |||||  || || | | || ||| |||
Db        1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy         987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
               | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db        1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy        1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                  |||||| || ||||||||| |||||  | |||||||| || |||   |   | || ||
Db        1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy        1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
                || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db        1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy        1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
               | ||    |    ||||| || ||||| || |||||||| | |||||  | ||||||||
Db        1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy        1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                 | || | |||||| || |||| ||  |||| |  ||||||   |  ||    || ||
Db        1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy        1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
               ||||||||||   ||| | |||||||||||||| |||||||||||  |  | || || |
Db        1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy        1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
               | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db        1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy        1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
               ||||||    |||| | | |  |||||  | ||| | | ||| ||| ||| |||||
Db        1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy        1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                 |  ||   | ||    ||||||| || ||||||  |  |   |||||    |    ||  |
Db        1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy        1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
               |||   |||||||||||||||| |||||||| || || ||||||||| ||    |||||| |
Db        1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy        1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                 |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db        1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy        1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                  ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db        1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy        1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
                |      | | ||| |        | || || |  || |  ||| ||||        ||
Db        1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy        1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
               ||  | || || ||| |||| | |       | |    ||  |  |    |||||||||
Db        2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy        1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
                | ||||||||| |||| |  |||||  | ||  | ||||||||||||||| ||    | |
Db        2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy        1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                   | ||| || | |   ||||| ||||||      ||||||  |||| ||||| |||
Db        2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy        1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
                |||  || ||||  |||||||   ||||||||||||       ||||||||||||||  |||
Db        2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy        2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
               | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db        2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy        2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
               |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db        2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy        2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
                |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db        2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy        2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
               || ||||  || | ||| ||||||||||||||||| |   ||  | || |    ||| ||
Db        2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy        2239 GCCGGGGCCCCGCCGCAGG 2257
                |||   |   |||   ||
Db        2527 CCCGCCCCGGAGCCCGCGG 2545


RESULT 13
AAH78204
ID   AAH78204 standard; DNA; 2316 BP.
XX
AC   AAH78204;
XX
DT   26-NOV-2001  (first entry)
XX
```

```
DE   Nucleotide sequence of a human secreted polypeptide.
XX
KW   Human; secreted polypeptide; nervous disease; muscular disease; tumour;
KW   gastrointestinal ulceration; spinal cord disease; trachea disease;
KW   thyroid gland disease; ovary disease; prostate disease; heart disease;
KW   renal gland disease; small intestine disease; thymus disease;
KW   lymph node disease; muscular system disease; colon disease;
KW   lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
KW   myocardial infarction; angioplasty; liver disease; coagulation disorder;
KW   microbial disease; immune disorder; inflammation; transplant rejection;
KW   bone thickness; bone density; ferroxidase loss; apoptosis;
KW   vascular smooth cell proliferation; vaccine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2316
FT                   /*tag=  a
FT                   /product= "secreted polypeptide"
XX
PN   WO200166690-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US07143.
XX
PR   06-MAR-2000; 2000US-0187107.
PR   13-MAR-2000; 2000US-0188916.
PR   03-OCT-2000; 2000US-0236874.
PR   03-OCT-2000; 2000US-0237846.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Agarwal P,  Murdoch PR,  Rizvi SK,  Smith RF,  Xiang Z;
XX
DR   WPI; 2001-570768/64.
DR   P-PSDB; AAG67512.
XX
PT   Novel isolated secreted polypeptide useful for treating nervous and
PT   muscular diseases, gastrointestinal ulceration, coagulation and immune
PT   disorders, microbial diseases, inflammation and transplant rejection -
XX
PS   Claim 2; Page 43-44; 102pp; English.
XX
CC   The present sequence encodes a human secreted polypeptide. The
CC   secreted polypeptides and polynucleotides are useful for treating
CC   nervous and muscular diseases, for inhibiting tumour formation and
CC   metastasis, for treating gastrointestinal ulceration, for preventing
CC   and treating diseases in spinal cord, thyroid gland, ovary, prostate,
CC   renal gland, small intestine, heart, trachea, thymus, lymph node,
CC   muscular system and colon, for treating lipase deficiency in cystic
CC   fibrosis and pancreatitis, for treating undesirable clot formation
CC   such as myocardial infarction, during angioplasty and all surgical
CC   procedures that require decreased blood clot formation, for treating
CC   liver diseases, coagulation disorders and microbial diseases, for
CC   treating immune disorders, for treating inflammation and transplant
CC   rejection, for enhancing bone thickness and increasing bone density,
CC   for reducing the loss of essential ferroxidases, for suppressing
CC   apoptosis, and for regulating vascular smooth cell proliferation. They
CC   may also be used as vaccines.
XX
SQ   Sequence 2316 BP; 347 A; 836 C; 758 G; 375 T; 0 other;

  Query Match             19.7%;  Score 618.8;  DB 22;  Length 2316;
  Best Local Similarity   60.7%;  Pred. No. 1e-120;
  Matches 1129;  Conservative    0;  Mismatches  677;  Indels   54;  Gaps    5;

Qy        487 CTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGC 546
              |||||   | |  |   || |    |    | | |  ||||||||     |||  |||||
Db         61 CTGCTGCTGCTCTGGGCGGGGGCATCTCGTGGCCAGCCCTGCCCCGGCCGCTGCATCTGC 120

Qy        547 CAGAATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCC 606
              |||||  || |     |||||   |  |||||| ||   |  ||  |||||||||| ||
Db        121 CAGAACGTGGCGCCCACACTGACAATGCTGTGCGCCAAGACCGGCTTGCTCTTTGTGCCG 180

Qy        607 CCTGATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATC 666
              || |  || ||||||||    |||||||||||| ||   || ||||||||||  |  | |
Db        181 CCCGCCATCGACCGGCGCGTGGTGGAGCTGCGGCTCACCGACAACTTCATCGCCGCCGTG 240

Qy        667 AGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACC 726
               ||||||  ||||| |||||||||||  | |||||| |||| || || ||| ||||||||
Db        241 CGCCGCCGAGACTTCGCCAACATGACCAGCCTGGTGCACCTCACTCTCTCCCGGAACACC 300

Qy        727 ATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTT 786
              ||| ||||  |     |     ||||    |||||      ||||||  ||||||| ||
Db        301 ATCGGCCAGGTGGCAGCTGGCGCCTTCGCCGACCTGCGTGCCCTCCGGGCCCTGCACCTG 360

Qy        787 GACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAG 846
              |||||||| || ||| |     |  | |  |||   ||||| ||||||| |||||| |
Db        361 GACAGCAACCGCCTGGCGGAGGTGCGCGGCGACCAGCTCCGCGGCCTGGGCAACCTCCGC 420

Qy        847 CACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTC 906
              ||||| ||| |    ||||||||| |  || |  | |     | ||| ||||| | ||||
Db        421 CACCTGATCCTTGGAAACAACCAGATCCGCCGGGTGGAGTCGGCGGCCTTTGACGCCTTC 480

Qy        907 CTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGAC 966
              |||   ||  ||||||| ||||| |||||||||||||  | | ||||||||||||
Db        481 CTGTCCACCGTGGAGGACCTGGATCTGTCCTACAACAACCTGGAGGCCCTGCCGTGGGAG 540

Qy        967 TCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCAC 1026
               | |||   |  ||||| ||||||  ||   || |  ||||||||||||||  | || |||
Db        541 GCGGTGGGCCAGATGGTGAACCTAAACACCCTCACGCTGGACCACAACCTCATCGACCAC 600

Qy       1027 ATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAAT 1086
              ||||| ||||| |||||     | || || || |||| ||| |||||  | ||||||||
Db        601 ATCGCGGAGGGGACCTTCGTGCAGCTTCACAAGCTGGTCCGTCTGGACATGACCTCCAAC 660

Qy       1087 CGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACA 1146
              || ||||| || || || || ||     ||||     | || ||||           ||
Db        661 CGCCTGCATAAACTCCCGCCCGACGGGCTCTTCCTGAGGTCGCAGG---------GCACC 711

Qy       1147 GCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAAT 1206
              |   |        || | ||  || || | || || || || ||||| || ||||||||
Db        712 GGGCCCAAGCCGCCCACCCCGCTGACCGTCAGCTTCGGCGGCAACCCCCTGCACTGCAAC 771

Qy       1207 TGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCC 1266
              || ||||| ||||||||||| ||||    || || || ||| | || ||||| | | |
Db        772 TGCGAGCTGCTCTGGCTGCGGCGGCTGACCCGCGAGGACGACTTAGAGACCTGCGCCACG 831

Qy       1267 CCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAG 1326
              || |    |||||  |  ||||||||||||    | |  |||||||||||  |||| |||
Db        832 CCCGAACACCTCACCGACCGCTACTTCTGGTCCATCCCCGAGGAGGAGTTCCTGTGTGAG 891

Qy       1327 CCGCCTCTCATCACCCAGCACAC---ACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCC 1383
              || || || ||||| | |||  |      |  |   |||||  |||| ||||||||||
Db        892 CCCCCGCTGATCACACGGCAGGCGGGGGGCCGGGCCCTGGTGGTGGAAGGCCAGGCGGTG 951

Qy       1384 ACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGAT 1443
              |  ||    |||  |||  | || ||||||   ||  |  | |||||||| || || |||
Db        952 AGCCTGCGCTGCCGAGCGGTGGGTGACCCCGAGCCGGTGGTGCACTGGGTGGCACCTGAT 1011

Qy       1444 GACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATC 1503
              |  || ||| | |||||||||    |||||   |||   |   | || || |||||  |
Db       1012 GGGCGGCTGCTGGGGAACTCCAGCCGGACCCGGGTCCGGGGGGACGGGACGCTGGATGTG 1071

Qy       1504 TTCATCACCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGA 1563
                 |||||||| |    |||||||||  ||||||| || || ||  |||||||||| ||
Db       1072 ACCATCACCACCTTGAGGGACAGTGGCACCTTCACTTGTATCGCCTCCAATGCTGCTGGG 1131

Qy       1564 GAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACC 1623
              || || |||||    |||||||| | | | || |  ||||| |        |      ||
Db       1132 GAAGCGACGGCGCCCGTGGAGGTGTGCGTGGTACCTCTGCCTC------TGATGGCACCC 1185

Qy       1624 AGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGC 1683
                 |   |||| || ||      |  |    | | |  | ||| |  |  || | |  ||
```

```
Db        1186 CCGCCGGCTGCCCCGCCGCCTCTCACCGAGCCCGGCTCCTCTGACATCGCCACGCCGGGC 1245

Qy        1684 CGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTG 1743
                |   |||||    |  |   |  |  |||| ||                        |
Db        1246 AGACCAGGTGCCAACGATTCTGCGGCTGAGCGTC---------------------GGCTC 1284

Qy        1744 CTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCA 1803
                | |   | ||  | ||| | | ||| |  ||  ||   ||| | | | |   | | | |
Db        1285 GTGGCAGCCGAGCTCACCTCGAACTCCGTGCTCATCCGCTGGCCAGCCCAGAGGCCTGTG 1344

Qy        1804 CCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTAC 1863
               ||| |  |    ||||||||| | |||||||||| | || |   |||||    ||  | |||
Db        1345 CCCGGAATACGCATGTACCAGGTTCAGTACAACAGTTCCGTTGATGACTCCCTCGTCTAC 1404

Qy        1864 AGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGC 1923
               |||||||||||  || |||   || ||||| |||| ||  |||||| | | ||   || |
Db        1405 AGGATGATCCCGTCCACCAGTCAGACCTTCCTGGTGAATGACCTGGCGGCGGGCCGTGCC 1464

Qy        1924 TACGACTTGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACCAAC 1983
               |||||||||||| ||||||||  | |  || |||   |||||  | ||  |||| ||
Db        1465 TACGACTTGTGCGTGCTGGCGGTCTACGACGACGGGGCCACAGCGCTGCCGGCAACGCGA 1524

Qy        1984 ATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCAC 2043
                | ||||||||| |  ||||||  ||||   |  ||     |||  ||||   |  ||
Db        1525 GTGGTGGGCTGTGTACAGTTCACCACCGCTGGGGATCCGGCGCCCTGCCGCCCGCTGAGG 1584

Qy        2044 AGCCAGATTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTG 2103
                 |||  |  ||||||||||||||||| | ||||||||||| ||||||| ||| || |
Db        1585 GCCCATTTCTTGGGCGGCACCATGATCATCGCCATCGGGGGCGTCATCGTCGCCTCGGTC 1644

Qy        2104 CTGGTCTTCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG- 2162
               || |||||||||||||  | |||||| | |||||| ||||| |    | ||| ||| |  | |
Db        1645 CTCGTCTTCATCGTTCTGCTCATGATCCGCTATAAGGTGTATGGCGACGGGGACAGCCGC 1704

Qy        2163 --------------CAAGATGGCAGCGGCCGTGAGCAATGTGTACTCGCAGACCAACGGC 2208
                             || |  |    ||    || ||| | |||| ||||||||||||||||
Db        1705 CGCGTCAAGGGCTCCAGGTCGCTCCCGCGGGTCAGCCACGTGTGCTCGCAGACCAACGGC 1764

Qy        2209 GCCCAGCCACCGCCTCCAAGCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAG 2268
               ||     ||    |  ||     | |  |      ||||| |   ||   |  | ||  |
Db        1765 GCAGGCACAGGCGCGGCACAGGCCCCGGCCCTGCCGGCCCAGGACCACTACGAGGCGCTG 1824

Qy        2269 GTGGTGGTGCGCAACGAGCTCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCC 2328
                  | ||||     | ||        |  |  ||||  |  ||||   ||||  ||  ||
Db        1825 CGCGAGGTGGAGTCCCAGGCTGCCCCCGCCGTCGCCGTCGAGGCCAAGGCCATGGAGGCC 1884
```

```
RESULT 14
AAD28124
ID   AAD28124 standard; DNA; 1653 BP.
XX
AC   AAD28124;
XX
DT   22-APR-2002  (first entry)
XX
DE   Human leucine-rich repeat-8 (ZLRR8) DNA #2.
XX
KW   Human; leucine-rich repeat-8; ZLRR8; cytostatic; gene therapy; leukaemia;
KW   endometrial adenocarcinoma; renal cell; colon; prostrate; retinoblastoma;
KW   cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;
KW   mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;
KW   rhabdomyosarcoma; genitourinary tract; chromosome 11q13; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1653
FT                   /*tag=  a
FT                   /product= "ZLRR8 protein"
FT                   /note= "CDS does not include stop codon"
FT                   /partial
FT   sig_peptide     1..54
FT                   /*tag=  b
FT   mat_peptide     55..1653
FT                   /*tag=  c
FT                   /product= "Mature ZLRR8 protein"
XX
PN   WO200202604-A2.
XX
PD   10-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-US20999.
XX
PR   30-JUN-2000; 2000US-215446P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Thayer EC,  Sheppard PO,  Presnell SR;
XX
DR   WPI; 2002-154725/20.
DR   P-PSDB; AAE17484.
XX
PT   New leucine-rich repeat proteins and polynucleotides, useful for
PT   diagnosing and treating disorders related to abnormal cell growth e.g.
PT   retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,
PT   kidney and lung tumors  -
XX
PS   Claim 19; Page 64-67; 82pp; English.
XX
CC   The invention relates to an isolated polypeptide comprising leucine-rich
CC   repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9.
CC   Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of
CC   proteins of interest from a host cell and to monitor the secretion of
CC   proteins in general from cells and tissues. The Zlrr DNA and proteins
CC   are useful in diagnosing and treating disorders related to abnormal cell
CC   growth, including retinoblastoma, renal cell adenocarcinoma, endometrial
CC   adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphotic leukaemia,
CC   kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,
CC   colon adenocarcinoma, genitourinary tract transitional cell tumours,
CC   rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas
CC   and prostate adenocarcinoma. Zlrr protein is useful for identifying
CC   agonists and antagonists of the polypeptide, for drug design, to screen
CC   for cell metabolism effecting receptors, for analysis of cell phenotype,
CC   and as animal feed supplement and cell culture components. Zlrr DNA is
CC   also useful in gene therapy. The present sequence is human Zlrr8 DNA.
CC   Zlrr8 gene is located on chromosome 11q13.
XX
SQ   Sequence 1653 BP; 223 A; 617 C; 526 G; 287 T; 0 other;

  Query Match             19.3%;  Score 608;  DB 24;  Length 1653;
  Best Local Similarity   63.7%;  Pred. No. 1.8e-118;
  Matches 1025;  Conservative    0;  Mismatches  540;  Indels   43;  Gaps    5;

Qy         520 GACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGC 579
               | |||||||||   |  ||| ||||||||||| ||||| ||||| ||  |  ||||| ||
Db          43 GCCGCCTGCCCGCTGCCCTGCGTCTGCCAGAACCTGTCCGAGTCGCTCAGCACCCTCTGT 102

Qy         580 CCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGC 639
                ||  |   || ||||| ||||| || ||  |  | |||||||| ||||||||||||||
Db         103 GCCCACCGAGGCCTGCTGTTTGTGCCGCCCAACGTGGACCGGCGCACAGTGGAGCTGCGG 162

Qy         640 CTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTG 699
               ||||  | ||||||||||     | |  | | ||  |||||   ||||||||||| |||
Db         163 CTGGCTGACAACTTCATCCAGGCCCTGGGGCCCCCTGACTTCCGCAACATGACGGGACTG 222

Qy         700 GTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGAC 759
               ||||||||||| |||||  | ||  ||||||| || |||   | ||    ||||  ||||
Db         223 GTGGACCTGACACTGTCTCGCAATGCCATCACCCGCATTGGGGCCCGCGCCTTTGGGGAC 282

Qy         760 CTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGAC 819
               ||||||||||| || ||||| || |||||| ||||  |||||      || ||    |
Db         283 CTCGAGAGCCTGCGTTCCCTCCACCTTGACGGCAACAGGCTGGTGGAGCTGGGCACCGGG 342

Qy         820 ACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGC 879
               | ||||||||||||  ||||| ||||||||||| ||| | | |  |||||||||||||| ||
Db         343 AGCCTCCGGGGCCCCGTCAATCTGCAGCACCTCATCCTCAGCGGCAACCAGCTGGGCCGC 402
```

```
Qy     880 ATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTAC 939
           |||||    |  || || || ||||||||   ||   ||||||| |||||||| ||||||
Db     403 ATCGCGCCGGGAGCCTTCGACGACTTCCTAGAGAGCCTGGAGGACCTGGACCTGTCCTAC 462

Qy     940 AACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTG 999
           ||||||||||      |||| |||| |  | |  |   ||||     ||| |||   ||
Db     463 AACAACCTCCGGCAGGTGCCCTGGGCCGGCATCGGCGCCATGCCTGCCCTGCACACCCTC 522

Qy    1000 AGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAA 1059
           | ||||||||| |||||  | ||     |  ||   ||| |||| ||  | ||     |
Db     523 AACCTGGACCATAACCTTATTGACGCACTGCCCCCAGGCGCCTTCGCCCAGCTCGGTCAG 582

Qy    1060 CTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTT 1119
           ||  |||||||||| ||||||||||| || |||   | |||| | || || ||  | ||
Db     583 CTCTCCCGCCTGGACCTCACCTCCAACCGCCTGGCCACGCTGGCTCCGGACCCGCTTTTC 642

Qy    1120 GCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGT 1179
            | ||    |  | |            ||     ||| | || | ||| ||    | ||
Db     643 TCTCGTGGGCGTGAT------------GCAGAGGCCTCTCCCGCCCCCCTGGTGCTGAGC 690

Qy    1180 TTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGG 1239
           ||| | || ||||| || |||||||| |||||||| || ||||||||| |||| | ||||
Db     691 TTTAGCGGGAACCCCCTGCACTGCAACTGTGAGCTGCTGTGGCTGCGGCGGCTGGCGCGG 750

Qy    1240 GACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCAT 1299
              || ||||||||||| || | ||||||    |||||    || ||||||||||||
Db     751 CCGGACGACCTGGAAACGTGCGCCTCCCCGCCCGGCCTGGCCGGCCGCTACTTCTGGGCA 810

Qy    1300 GTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTG 1359
           ||||  ||||  |||||    || |||||||| |||||  |||  ||||| ||     |
Db     811 GTGCCCGAGGGCGAGTTCTCCTGTGAGCCGCCCCTCATTGCCCGCCACACGCAGCGCCTC 870

Qy    1360 CTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCC 1419
             ||| ||||| ||||||  |||||| ||   ||||   ||| | || ||||||   ||
Db     871 TGGGTGCTGGAAGGCCAGCGGGCCACGCTGCGGTGCCGGGCCCTGGGTGACCCCGCGCCT 930

Qy    1420 CTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTC 1479
              || |||||||| |  || || |||||  |||| || ||||||||  |  ||   |
Db     931 ACCATGCACTGGGTCGGTCCTGACGACCGGTTGGTTGGCAACTCCTCCCGAGCCCGGGCT 990

Qy    1480 TATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 1539
           |    ||| || ||| | || ||     | |||    ||  ||||  ||| | || ||||
Db     991 TTCCCCAACGGGACCTTAGAGATTGGGGTGACCGGCGCTGGGGACGCTGGGGGCTACACC 1050

Qy    1540 TGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAG 1599
           ||||| ||  ||||  |||| || |||||||| |||   || ||  |     |  |
Db    1051 TGCATCGCCACCAACCCTGCTGGTGAGGCCACAGCCCGAGTAGAACTGCGGGTGCTGGCC 1110

Qy    1600 CTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGAC 1659
            |||| ||     | |||||||    |      |  | ||      || | | ||| |||
Db    1111 TTGCCCCATGGTGGGAACAGCAGTGCCGAGGGGGGCCGCC------CCGGGCCCTCGGAC 1164

Qy    1660 ATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCC 1719
           ||| | | ||||      ||       |  ||||  || |    ||     |||||
Db    1165 ATCGCCGCCTCCGCTCGCACTGCTGCCGAGGGTGAGGGGACGCTGGAGTCTGAGCC---- 1220

Qy    1720 AAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTC 1779
                             || ||||  ||| | || |||||| ||||||| |  |||||
Db    1221 -----------------AGCCGTGCAGGTGACGGAGGTGACCGCCACCTCAGGGCTGGTG 1263

Qy    1780 AAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGTACAACTGC 1839
           |  |||  |  | |   | ||||   ||  |||  ||||| |||  | ||||||||| ||
Db    1264 AGCTGGGGTCCCGGGCGGCCAGCCGACCCAGTGTGGATGTTCCAAATCCAGTACAACAGC 1323

Qy    1840 TCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCTTCGTGGTC 1899
              || ||||||   || || ||| ||||  |||||||||||| | |   |||| || |
Db    1324 AGCGAAGATGAGACCCTCATCTACCGGATTGTCCCAGCCTCCAGCCACCACTTCCTGCTG 1383

Qy    1900 AACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGGATGACACA 1959
           ||  |||| ||  | ||  ||| ||| ||| | ||  |||||||| |||     | | |
Db    1384 AAGCACCTCGTCCCCGGCGCTGACTATGACCTCTGCCTGCTGGCCTTGTCACCGGCCGCT 1443
```

```
Qy    1960 G---CCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCACCAAGGCT 2016
           |   || |    |||||||||||||   |  ||||||| ||||| |||| |||   | |
Db    1444 GGGCCCTCTGACCTCACGGCCACCAGGCTGCTGGGCTGTGCCCATTTCTCCACGCTGCCG 1503

Qy    2017 GACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGATCCTGGTC 2076
           | ||  || | ||||||  || ||||   |||  | |||||||| ||| ||| | ||| |
Db    1504 GCCTCGCCCCTGTGCCACGCCCTGCAGGCCCACGTGCTGGGCGGGACCCTGACCGTGGCC 1563

Qy    2077 ATCGGGGGCATCATCGTGGCCACGC-TGCTGGTCTTCATCGTCATCCT 2123
            | |||||  |  | |||||  ||| | ||||||||||  ||   |||
Db    1564 GTGGGGGGTGTGCTGGTGGCTGCGCTTACTGGTCTTCACTGTGGCCCT 1611

RESULT 15
AAD28125
ID   AAD28125 standard; DNA; 1653 BP.
XX
AC   AAD28125;
XX
DT   22-APR-2002  (first entry)
XX
DE   Human leucine-rich repeat-8 (ZLRR8) degenerate DNA #2.
XX
KW   Human; leucine-rich repeat-8; ZLRR8; cytostatic; gene therapy; leukaemia;
KW   endometrial adenocarcinoma; renal cell; colon; prostrate; retinoblastoma;
KW   cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;
KW   mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;
KW   rhabdomyosarcoma; genitourinary tract; chromosome 11q13; ds.
XX
OS   Homo sapiens.
XX
PN   WO200202604-A2.
XX
PD   10-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-US20999.
XX
PR   30-JUN-2000; 2000US-215446P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Thayer EC,  Sheppard PO,  Presnell SR;
XX
DR   WPI; 2002-154725/20.
XX
PT   New leucine-rich repeat proteins and polynucleotides, useful for
PT   diagnosing and treating disorders related to abnormal cell growth e.g.
PT   retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,
PT   kidney and lung tumors  -
XX
PS   Disclosure; Page 69-70; 82pp; English.
XX
CC   The invention relates to an isolated polypeptide comprising leucine-rich
CC   repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9.
CC   Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of
CC   proteins of interest from a host cell and to monitor the secretion of
CC   proteins in general from cells and tissues. The Zlrr DNA and proteins
CC   are useful in diagnosing and treating disorders related to abnormal cell
CC   growth, including retinoblastoma, renal cell adenocarcinoma, endometrial
CC   adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphotic leukaemia,
CC   kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,
CC   colon adenocarcinoma, genitourinary tract transitional cell tumours,
CC   rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas
CC   and prostate adenocarcinoma. Zlrr protein is useful for identifying
CC   agonists and antagonists of the polypeptide, for drug design, to screen
CC   for cell metabolism effecting receptors, for analysis of cell phenotype,
CC   and as animal feed supplement and cell culture components. Zlrr DNA is
CC   also useful in gene therapy. The present sequence is human Zlrr8
CC   degenerate DNA. Zlrr8 gene is located on chromosome 11q13.
XX
SQ   Sequence 1653 BP; 179 A; 217 C; 329 G; 195 T; 733 other;

  Query Match             17.6%;  Score 554.6;  DB 24;  Length 1653;
```

```
Best Local Similarity   40.1%;  Pred. No. 3.2e-107;
Matches  643;  Conservative  276;  Mismatches  642;  Indels   42;  Gaps    4;

Qy         512 CCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGA 571
               : |  |  |  || ||:||       ||:|| ||:||:||::| :: ||::: :|  : |
Db          35 SNGGNGCNGCNGCNTGYCCNYTNCCNTGYGTNTGYCARAAYYTNWSNGARWSNYTNWSNA 94

Qy         572 CCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGG 631
               | :| ||: |   ::  || :| :| ||:|| || ||  |: | ||::| :| || || |
Db          95 CNYTNTGYGCNCAYMGNGGNYTNYTNTTYGTNCCNCCNAAYGTNGAYMGNMGNACNGTNG 154

Qy         632 AGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGA 691
               |::| :| :| |  | :||:||:||:       |  | |  |  ||:||:   ||:||||
Db         155 ARYTNMGNYTNGCNGAYAAYTTYATHCARGCNYTNGGNCCNCCNGAYTTYMGNAAYATGA 214

Qy         692 CGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCT 751
               | || :| || ||::| || :| :: :| ||: | ||:|  :  ||:    |     | |
Db         215 CNGGNYTNGTNGAYYTNACNYTNWSNMGNAAYGCNATHACNMGNATHGGNGCNMGNGCNT 274

Qy         752 TTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTG 811
               |:   ||::| ||::: :| :| :: :| ||::| ||: | ||::| :|       :| |
Db         275 TYGGNGAYYTNGARWSNYTNMGNWSNYTNCAYYTNGAYGGNAAYMGNYTNGTNGARYTNG 334

Qy         812 GGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGC 871
               |    |  :: :| :| || |  || ||::| ||:||::| ||: | :     ||:||::
Db         335 GNACNGGNWSNYTNMGNGGNCCNGTNAAYYTNCARCAYYTNATHYTNWSNGGNAAYCARY 394

Qy         872 TGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACC 931
               | ||  | ||:||    |  || ||:|| ||:||::|   ::: :| ||:||::| ||::
Db         395 TNGGNMGNATHGCNCCNGGNGCNTTYGAYGAYTTYYTNGARWSNYTNGARGAYYTNGAYY 454

Qy         932 TCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCC 991
               | :: ||:||:||::| :      | || ||||      |  |    |||      :| |
Db         455 TNWSNTAYAAYAAYYTNMGNCARGTNCCNTGGGCNGGNATHGGNGCNATGCCNGCNYTNC 514

Qy         992 ACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACC 1051
               |:   :| | ::| ||:||:||::|  | ||:    |  |    ||  | ||:||  | :
Db         515 AYACNYTNAAYYTNGAYCAYAAYYTNATHGAYGCNYTNCCNCCNGGNGCNTTYGCNCARY 574

Qy        1052 TGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATC 1111
               |     |::|  : :| :| ||::| || :: ||::| :|    |  :|  | || ||:|
Db         575 TNGGNCARYTNWSNMGNYTNGAYYTNACNWSNAAYMGNYTNGCNACNYTNGCNCCNGAYC 634

Qy        1112 CCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGT 1171
               |  | ||: : :|    :  | :            ||   : | :   |  | || :|
Db         635 CNYTNTTYWSNMGNGGNMGNGAY------------GCNGARGCNWSNCCNGCNCCNYTNG 682

Qy        1172 CCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGC 1231
                 :| :: ||: : || ||:|| :| ||:||:||:||:||::| :| |||:| :| :| :
Db         683 TNYTNWSNTTYWSNGGNAAYCCNYTNCAYTGYAAYTGYGARYTNYTNTGGYTNMGNMGNY 742

Qy        1232 TCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACT 1291
               | |  :|    ||:||::| ||:|| ||:|  :: ||    || :|    || :| ||:|
Db         743 TNGCNMGNCCNGAYGAYYTNGARACNTGYGCNWSNCCNCCNGGNYTNGCNGGNMGNTAYT 802

Qy        1292 TCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACAC 1351
               |:|||   || |  ||:|  ||:||:   ||:||:|| || :| ||: | :  ||:|| |
Db         803 TYTGGGCNGTNCCNGARGGNGARTTYWSNTGYGARCCNCCNYTNATHGCNMGNCAYACNC 862

Qy        1352 ACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACC 1411
               | :  :|   ||| :| ||:|| ||:   || || :| :  ||::  ||  | || ||:|
Db         863 ARMGNYTNTGGGTNYTNGARGGNCARMGNGCNACNYTNMGNTGYMGNGCNYTNGGNGAYC 922

Qy        1412 CCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGA 1471
               |    ||    || ||:||||| |  || ||:||::| :| || || ||::: :: :|
Db         923 CNGCNCCNACNATGCAYTGGGTNGGNCCNGAYGAYMGNYTNGTNGGNAAYWSNWSNMGNG 982

Qy        1472 CCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTG 1531
               |    |  | :   ||:|| || :| || ||:    | ||     |    ||:   || |
Db         983 CNMGNGCNTTYCCNAAYGGNACNYTNGARATHGGNGTNACNGGNGCNGGNGAYGCNGGNG 1042

Qy        1532 CCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCA 1591
                 | :|| ||:||:||  | ||: | || || ||:|| || || :  || ||: |
Db        1043 GNTAYACNTGYATHGCNACNAAYCCNGCNGGNGARGCNACNGCNMGNGTNGARYTNMGNG 1102

Qy        1592 TCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCC 1651
               |  |    :| || ||:    | ||::: ::        | | :      :   |  | |
Db        1103 TNYTNGCNYTNCCNCAYGGNGGNAAYWSNWS------NGCNGARGGNGGNMGNCCNGGNC 1156

Qy        1652 TCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAG 1711
                 :: ||:||: | |  ::    :  ||       | :|| | :|| |     |   : |
Db        1157 CNWSNGAYATHGCNGCNWSNGCNMGNACNGCNGCNGARGGNGARGGNACNYTNGARWSNG 1216

Qy        1712 AGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGG 1771
               |:||                      || || |  ||  | ||:|| ||  | || :: |
Db        1217 ARCC---------------------NGCNGTNCARGTNACNGARGTNACNGCNACNWSNG 1255

Qy        1772 CCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGT 1831
                 :| || :  |||        | :   | ||   :|  ||   ||||| :||: | ||:|
Db        1256 GNYTNGTNWSNTGGGGNCCNGGNMGNCCNGCNGAYCCNGTNTGGATGTTYCARATHCART 1315

Qy        1832 ACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAGGCCT 1891
               |:||::: :: || ||:||:   :| ||:||::| || ||  || || :: :   |   :|
Db        1316 AYAAYWSNWSNGARGAYGARACNYTNATHTAYMGNATHGTNCCNGCNWSNWSNCAYCAYT 1375

Qy        1892 TCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGG 1951
               |: |  | ||  |::| ||  | ||  | | :||:||::| ||: | :| ||  | ::
Db        1376 TYYTNYTNAARCAYYTNGTNCCNGGNGCNGAYTAYGAYYTNTGYYTNYTNGCNYTNWSNC 1435

Qy        1952 ATGACACAGCCACGACA---CTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTCTTCA 2008
                 |   | |   | ::    :| || || || :   |  |  || ||:|| || ||::  |
Db        1436 CNGCNGCNGGNCCNWSNGAYYTNACNGCNACNMGNYTNYTNGGNTGYGCNCAYTTYWSNA 1495

Qy        2009 CCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACCATGA 2068
               |     | |  :  || :  ||:||  |  | ||    ||  | :| || || ||  | |
Db        1496 CNYTNCCNGCNWSNCCNYTNTGYCAYGCNYTNCARGCNCAYGTNYTNGGNGGNACNYTNA 1555

Qy        2069 TCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTT 2111
                  | |   | || ||     | || ||  | :|    |  :|
Db        1556 CNGTNGCNGTNGGNGGNGTNYTNGTNGCNGCNYTNACNGGNYT 1598

Search completed: June 8, 2003, 15:46:32
Job time : 698 secs
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2003  Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:          June  8, 2003, 15:34:52 ; Search time 175 Seconds
                                            (without alignments)
                                            5509.667 Million cell updates/sec

Title:           US-09-831-846-1
Perfect score:   3144
Sequence:        1 gcctggctccctctcgctga..........gcggctttggattgcttatg 3144

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        441362 seqs, 153338381 residues .

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                     %
Result               Query
  No.     Score   Match Length DB  ID                         Description
-------------------------------------------------------------------------------
     1     81.8     2.6    7452  3  US-08-592-500-1              Sequence 1, Appli
     2     81.8     2.6    7452  3  US-08-195-006-1              Sequence 1, Appli
     3     81.8     2.6    7452  5  PCT-US94-07644A-1            Sequence 1, Appli
     4     81.6     2.6    2852  4  US-09-063-950-1              Sequence 1, Appli
     5     79.4     2.5    2019  4  US-09-063-950-3              Sequence 3, Appli
     6     77.2     2.5    5176  4  US-09-182-024A-1             Sequence 1, Appli
     7     63.8     2.0    4843  3  US-08-986-485-1              Sequence 1, Appli
     8     56.4     1.8    2600  2  US-08-427-497E-4             Sequence 4, Appli
     9     56.4     1.8    3189  2  US-08-427-497E-3             Sequence 3, Appli
    10     56.4     1.8    3774  2  US-08-341-843B-1             Sequence 1, Appli
    11     56.4     1.8    3774  2  US-08-427-497E-1             Sequence 1, Appli
    12     56.4     1.8    3774  2  US-08-427-497E-2             Sequence 2, Appli
    13     56.4     1.8    3888  4  US-08-506-296B-13            Sequence 13, Appl
    14     51.6     1.6    4203  2  US-08-866-757-1              Sequence 1, Appli
    15     51.6     1.6    4203  4  US-09-153-593-1              Sequence 1, Appli
    16     50.2     1.6    1685  6  5340934-1                   Patent No. 5340934
c   17     48.8     1.6    1362  4  US-08-818-112-7              Sequence 7, Appli
c   18     48.8     1.6    1362  4  US-08-818-111-7              Sequence 7, Appli
c   19     48.8     1.6    1362  4  US-09-056-556-7              Sequence 7, Appli
c   20     48.8     1.6    1362  4  US-09-072-596-7              Sequence 7, Appli
    21     48.8     1.6 4403765  4  US-09-103-840A-2              Sequence 2, Appli
    22     48.8     1.6 4411529  4  US-09-103-840A-1              Sequence 1, Appli
c   23       48     1.5     503  4  US-09-297-269-40             Sequence 40, Appl
    24     47.6     1.5    3783  4  US-08-506-296B-20            Sequence 20, Appl
    25     47.4     1.5     906  1  US-07-817-852A-1             Sequence 1, Appli
    26     47.4     1.5     906  1  US-08-470-137-1              Sequence 1, Appli
    27     47.4     1.5    1182  4  US-09-385-028-19             Sequence 19, Appl
```

```
    28     47.4     1.5   11604  4  US-09-385-028-13             Sequence 13, Appl
    29     47.4     1.5   15079  4  US-09-385-028-1              Sequence 1, Appli
c   30     47.2     1.5    2338  1  US-08-425-069-1              Sequence 1, Appli
c   31     47.2     1.5    2338  2  US-08-317-844B-1             Sequence 1, Appli
    32     46.6     1.5   44377  2  US-08-804-227C-7             Sequence 7, Appli
    33     46.6     1.5   44377  2  US-08-804-198-1              Sequence 1, Appli
    34     46.4     1.5    8378  5  PCT-US91-09055-1             Sequence 1, Appli
    35     46.2     1.5    4134  4  US-09-162-021B-1             Sequence 1, Appli
    36     46.2     1.5    4134  4  US-09-687-477-17             Sequence 17, Appl
    37     45.8     1.5    2290  4  US-09-131-648-4              Sequence 4, Appli
    38     45.4     1.4     891  4  US-09-452-239-1              Sequence 1, Appli
c   39     45.4     1.4 4403765  4  US-09-103-840A-2              Sequence 2, Appli
    40     45.2     1.4    1018  4  US-09-452-239-35             Sequence 35, Appl
    41     44.8     1.4    1119  4  US-08-987-943-2              Sequence 2, Appli
    42     44.8     1.4    1755  4  US-08-987-943-1              Sequence 1, Appli
c   43     44.8     1.4    1926  4  US-09-249-585A-2             Sequence 2, Appli
c   44     44.8     1.4    2580  3  US-09-050-863-2              Sequence 2, Appli
c   45     44.8     1.4    2580  4  US-09-359-081-2              Sequence 2, Appli


                                 ALIGNMENTS


RESULT 1
US-08-592-500-1
; Sequence 1, Application US/08592500
; Patent No. 6005089
;  GENERAL INFORMATION:
;    APPLICANT:  Lanza, Francois
;    APPLICANT:  Phillips, David R.
;    APPLICANT:  Cazenave, Jean-Pierre
;    TITLE OF INVENTION:  Platelet Glycoprotein V Gene and Uses
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,500
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/089,455
;      FILING DATE:  09-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Karen B.
;      REGISTRATION NUMBER:  29,684
;      REFERENCE/DOCKET NUMBER:  12418-28
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7452 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  1462..2419
;    FEATURE:
;      NAME/KEY:  CDS
```

```
;      LOCATION:  2422..4101
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  68..76
;      OTHER INFORMATION:  /function= "Putitive TPA responsive
;      OTHER INFORMATION:  element"
;      OTHER INFORMATION:  /label= TRE
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  471..478
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  493..502
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  593..881
;      OTHER INFORMATION:  /rpt_type= "other"
;      OTHER INFORMATION:  /label= Alu
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  897..904
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1142..1149
;      OTHER INFORMATION:  /function= "Sp1 binding site"
;      OTHER INFORMATION:  /standard_name= "Sp1"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  1178..1184
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequences"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  TATA_signal
;      LOCATION:  1199..1203
;    FEATURE:
;      NAME/KEY:  TATA_signal
;      LOCATION:  1263..1269
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1285..1289
;      OTHER INFORMATION:  /function= "GATA-1 binding site"
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1321..1326
;      OTHER INFORMATION:  /function= "GATA-1 binding site"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  1365..1372
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequences"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  6133..6440
;      OTHER INFORMATION:  /rpt_type= "other"
;      OTHER INFORMATION:  /label= Alu
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  5610..5615
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
```

```
;      LOCATION:  6966..6971
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7224..7229
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7358..7363
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..7452
;      OTHER INFORMATION:  /standard_name= "Nucleotide
;      OTHER INFORMATION:  sequence containing the human GPV gene"
US-08-592-500-1

  Query Match             2.6%;  Score 81.8;  DB 3;  Length 7452;
  Best Local Similarity   50.4%;  Pred. No. 1.1e-08;
  Matches  200;  Conservative    0;  Mismatches  197;  Indels    0;  Gaps    0;

Qy        707 TGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGA 766
              |||| |||| |  |||| |  |     |  ||| |    |   |||   |||  | | |
Db       3233 TGACTCTGTTCGAGAACCCGCTGGCAGAGCTCCCGGGGGTGCTCTTCGGGGAGATGGGGG 3292

Qy        767 GCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCC 826
              |||| |     |||   ||  || |||  | |||||  | |||    |  | | || |||
Db       3293 GCCTGCAGGAGCTGTGGCTGAACCGCACCCAGCTGCGCACCCTGCCCGCCGCCGCCTTCC 3352

Qy        827 GGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAG 886
              |   ||||  |  ||||| | || |    ||||     | || || |    ||
Db       3353 GCAACCTGAGCCGCCTGCGGTACTTAGGGGTGACTCTGAGCCCGCGGCTGAGCGCGCTTC 3412

Qy        887 ATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACC 946
                 |||         |   |||        |  |||  || | |||   || ||||  ||
Db       3413 CGCAGGGCGCCTTCCAGGGCCTTGGCGAGCTCCAGGTGCTCGCCCTGCACTCCAACGGCC 3472

Qy        947 TCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGG 1006
              |    | ||| ||    | ||   ||||  || | | ||| || | |||| ||  ||||
Db       3473 TGACCGCCCTCCCCGACGGCTTGCTGCGCGGCCTCGGCAAGCTGCGCCAGGTGTCCCTGC 3532

Qy       1007 ACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCC 1066
               || ||||  ||||     | |  ||   | |  |||     | ||    |  ||||
Db       3533 GCCGCAACAGGCTGCGCGCCCTGCCCCGTGCCCTCTTCCGCAATCTCAGCAGCCTGGAGA 3592

Qy       1067 GCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCC 1103
              || |  | |||  |  ||| | |||| |||  |||||
Db       3593 GCGTCCAGCTCGACCACAACCAGCTGGAGACCCTGCC 3629


RESULT 2
US-08-195-006-1
; Sequence 1, Application US/08195006
; Patent No. 6083688
;  GENERAL INFORMATION:
;    APPLICANT:  Lanza, Francois
;    APPLICANT:  Phillips, David R.
;    APPLICANT:  Cazenave, Jean-Pierre
;    TITLE OF INVENTION:  Platelet Glycoprotein V Gene and Uses
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/195,006
;      FILING DATE:  10-FEB-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/089,455
;      FILING DATE:  09-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Karen B.
;      REGISTRATION NUMBER:  29,684
;      REFERENCE/DOCKET NUMBER:  12418-28
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7452 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  1462..2419
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2422..4101
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  68..76
;      OTHER INFORMATION:  /function= "Putitive TPA responsive
;      OTHER INFORMATION:  element"
;      OTHER INFORMATION:  /label= TRE
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  471..478
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  493..502
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  593..881
;      OTHER INFORMATION:  /rpt_type= "other"
;      OTHER INFORMATION:  /label= Alu
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  897..904
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1142..1149
;      OTHER INFORMATION:  /function= "Sp1 binding site"
;      OTHER INFORMATION:  /standard_name= "Sp1"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  1178..1184
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequences"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  TATA_signal
;      LOCATION:  1199..1203
;    FEATURE:
;      NAME/KEY:  TATA_signal
;      LOCATION:  1263..1269
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1285..1289
;      OTHER INFORMATION:  /function= "GATA-1 binding site"
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1321..1326
;      OTHER INFORMATION:  /function= "GATA-1 binding site"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  1365..1372
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequences"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  6133..6440
;      OTHER INFORMATION:  /rpt_type= "other"
;      OTHER INFORMATION:  /label= Alu
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  5610..5615
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  6966..6971
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7224..7229
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7358..7363
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..7452
;      OTHER INFORMATION:  /standard_name= "Nucleotide
;      OTHER INFORMATION:  sequence containing the human GPV gene"
US-08-195-006-1

  Query Match             2.6%;  Score 81.8;  DB 3;  Length 7452;
  Best Local Similarity   50.4%;  Pred. No. 1.1e-08;
  Matches  200;  Conservative    0;  Mismatches  197;  Indels    0;  Gaps    0;

Qy        707 TGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGA 766
              |||| |||| |  |||| |  |     |  ||| |    |   |||   |||  | | |
Db       3233 TGACTCTGTTCGAGAACCCGCTGGCAGAGCTCCCGGGGGTGCTCTTCGGGGAGATGGGGG 3292

Qy        767 GCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCC 826
              |||| |     |||   ||  || |||  | |||||  | |||    |  | | || |||
Db       3293 GCCTGCAGGAGCTGTGGCTGAACCGCACCCAGCTGCGCACCCTGCCCGCCGCCGCCTTCC 3352

Qy        827 GGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAG 886
              |   ||||  |  ||||| | || |    ||||     | || || |    ||
Db       3353 GCAACCTGAGCCGCCTGCGGTACTTAGGGGTGACTCTGAGCCCGCGGCTGAGCGCGCTTC 3412

Qy        887 ATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACC 946
                 |||         |   |||        |  |||  || | |||   || ||||  ||
Db       3413 CGCAGGGCGCCTTCCAGGGCCTTGGCGAGCTCCAGGTGCTCGCCCTGCACTCCAACGGCC 3472

Qy        947 TCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGG 1006
              |    | ||| ||    | ||   ||||  || | | ||| || | |||| ||  ||||
Db       3473 TGACCGCCCTCCCCGACGGCTTGCTGCGCGGCCTCGGCAAGCTGCGCCAGGTGTCCCTGC 3532
```

```
Qy       1007 ACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCC 1066
               || ||||  ||||     | |  ||   | |  |||     | ||    |  ||||
Db       3533 GCCGCAACAGGCTGCGCGCCCTGCCCCGTGCCCTCTTCCGCAATCTCAGCAGCCTGGAGA 3592

Qy       1067 GCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCC 1103
              || |  | |||  |  ||| | |||| |||  |||||
Db       3593 GCGTCCAGCTCGACCACAACCAGCTGGAGACCCTGCC 3629


RESULT 3
PCT-US94-07644A-1
; Sequence 1, Application PC/TUS9407644A
;  GENERAL INFORMATION:
;    APPLICANT:  COR Therapeutics, Inc.
;    TITLE OF INVENTION:  Platelet Glycoprotein V Gene and Uses
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/07644A
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Karen B.
;      REGISTRATION NUMBER:  29,684
;      REFERENCE/DOCKET NUMBER:  012418-003000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7452 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  1462..2419
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2422..4101
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  68..76
;      OTHER INFORMATION:  /function= "Putitive TPA responsive
;      OTHER INFORMATION:  element"
;      OTHER INFORMATION:  /label= TRE
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  471..478
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  493..502
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  593..881
;      OTHER INFORMATION:  /rpt_type= "other"
;      OTHER INFORMATION:  /label= Alu
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  897..904
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequence"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1142..1149
;      OTHER INFORMATION:  /function= "Sp1 binding site"
;      OTHER INFORMATION:  /standard_name= "Sp1"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  1178..1184
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequences"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  TATA_signal
;      LOCATION:  1199..1203
;    FEATURE:
;      NAME/KEY:  TATA_signal
;      LOCATION:  1263..1269
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1285..1289
;      OTHER INFORMATION:  /function= "GATA-1 binding site"
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  1321..1326
;      OTHER INFORMATION:  /function= "GATA-1 binding site"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  1365..1372
;      OTHER INFORMATION:  /function= "Ets-1 cis-acting
;      OTHER INFORMATION:  sequences"
;      OTHER INFORMATION:  /label= Ets-1
;    FEATURE:
;      NAME/KEY:  repeat_region
;      LOCATION:  6133..6440
;      OTHER INFORMATION:  /rpt_type= "other"
;      OTHER INFORMATION:  /label= Alu
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  5610..5615
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  6966..6971
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7224..7229
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7358..7363
;      OTHER INFORMATION:  /standard_name= "Polyadenylation
;      OTHER INFORMATION:  signal sequence"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..7452
;      OTHER INFORMATION:  /standard_name= "Nucleotide
;      OTHER INFORMATION:  sequence containing the human GPV gene"
PCT-US94-07644A-1
```

```
  Query Match             2.6%;  Score 81.8;  DB 5;  Length 7452;
  Best Local Similarity   50.4%;  Pred. No. 1.1e-08;
  Matches  200;  Conservative    0;  Mismatches  197;  Indels    0;  Gaps    0;

Qy        707 TGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGA 766
              |||| |||| |  |||| |  |     |  ||| |    |   |||   |||  | | |
Db       3233 TGACTCTGTTCGAGAACCCGCTGGCAGAGCTCCCGGGGGTGCTCTTCGGGGAGATGGGGG 3292

Qy        767 GCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCC 826
              |||| |     |||   ||  || |||  | |||||  | |||    |  | | || |||
Db       3293 GCCTGCAGGAGCTGTGGCTGAACCGCACCCAGCTGCGCACCCTGCCCGCCGCCGCCTTCC 3352

Qy        827 GGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAG 886
              |   ||||  |  ||||| | || |    ||||     | || || |    ||
Db       3353 GCAACCTGAGCCGCCTGCGGTACTTAGGGGTGACTCTGAGCCCGCGGCTGAGCGCGCTTC 3412

Qy        887 ATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACC 946
                 |||         |   |||        |  |||  || | |||   || ||||  ||
Db       3413 CGCAGGGCGCCTTCCAGGGCCTTGGCGAGCTCCAGGTGCTCGCCCTGCACTCCAACGGCC 3472

Qy        947 TCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGG 1006
              |    | ||| ||    | ||   ||||  || | | ||| || | |||| ||  ||||
Db       3473 TGACCGCCCTCCCCGACGGCTTGCTGCGCGGCCTCGGCAAGCTGCGCCAGGTGTCCCTGC 3532

Qy       1007 ACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCC 1066
               || ||||  ||||     | |  ||   | |  |||     | ||    |  ||||
Db       3533 GCCGCAACAGGCTGCGCGCCCTGCCCCGTGCCCTCTTCCGCAATCTCAGCAGCCTGGAGA 3592

Qy       1067 GCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCC 1103
              || |  | |||  |  ||| | |||| |||  |||||
Db       3593 GCGTCCAGCTCGACCACAACCAGCTGGAGACCCTGCC 3629

RESULT 4
US-09-063-950-1
; Sequence 1, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION: THEREFOR
;  FILE REFERENCE: MEI-019
;  CURRENT APPLICATION NUMBER: US/09/063,950C
;  CURRENT FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (160)..(2178)
US-09-063-950-1

  Query Match             2.6%;  Score 81.6;  DB 4;  Length 2852;
  Best Local Similarity   46.6%;  Pred. No. 8.9e-09;
  Matches  338;  Conservative    0;  Mismatches  379;  Indels    9;  Gaps    2;

Qy        315 CCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAAGGATGCGTGGCCGA 374
              ||||||||   || ||||  |    || ||   | | |  |     |||  |  | ||
Db         15 CCGGAGCCCGGGGCGGGTGGACGCGGACTCGAACGCAGTTGCTTCGGGACCCAGGACCCC 74

Qy        375 GCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCAGCCGGGGGTAGATG 434
                ||||     |  ||       |     |||| |||     |||  ||||        |
Db         75 CTCGGGCCCGACCCGCCAGGAAAGACTGAGGCCGCGGCCTGCCCCGCCCGGCTCCCTGCG 134

Qy        435 CTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGGCCTGCTAGC 494
              | ||| |  ||   ||      |  ||||  |||    ||  | |   |   |||||  |
Db        135 CCGCCGCCGCCTCCCG------GGACAGAAGATGTGCTCCAGGGTCCCTCTGCTGCTGCC 188

Qy        495 GTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCT 554
              | |   | ||    | ||| |||     |     |     | |     | ||    |
Db        189 GCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGGGCTGCCCATCCGGCTGCCAGTG 248

Qy        555 GTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATAT 614
                   || ||| |     | | ||| |  ||     |  |  |   || ||||  ||  |
Db        249 CAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCCCGAGACGT 308

Qy        615 TGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCA 674
                  |     || |||| ||||  | |    |  |||  ||||| |    ||  |
Db        309 GCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGCAGG 368

Qy        675 GGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCA 734
                 |||||||  | || |||| |||  |  ||||  ||||||   ||||   |||  |
Db        369 CAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCACAGAACCAGATCGCCAG 428

Qy        735 CATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAA 794
              | | |    |     ||| | |   ||||  | |||| ||  |||| | ||      |||
Db        429 CCTGCCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAA 488

Qy        795 TCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTAT 854
                ||||||      |     | || ||| |||| ||||||     |||  ||| |||
Db        489 CAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTA 548

Qy        855 CGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGAC 914
              | ||  ||| ||||   |  ||  ||||     ||  || || ||     ||   |
Db        549 CCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCT 608

Qy        915 ATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCG 974
                |||||   ||  | ||     ||||| | || |  |  |||||   |     | |||
Db        609 CCTGGAG---CTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGCCTGCC 665

Qy        975 ACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGA 1034
               ||| || |    ||  |||       |     || |      ||||| | |  |  |
Db        666 CCGCCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCT 725

Qy       1035 GGGCAC 1040
              || |||
Db        726 GGACAC 731

RESULT 5
US-09-063-950-3
; Sequence 3, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION: THEREFOR
;  FILE REFERENCE: MEI-019
;  CURRENT APPLICATION NUMBER: US/09/063,950C
;  CURRENT FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2019
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2019)
US-09-063-950-3

  Query Match             2.5%;  Score 79.4;  DB 4;  Length 2019;
  Best Local Similarity   49.3%;  Pred. No. 2.4e-08;
  Matches  237;  Conservative    0;  Mismatches  241;  Indels    3;  Gaps    1;

Qy        560 AGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACC 619
              || ||| |     | | ||| |  ||      |  |  |   || ||||  ||  |    |
Db         95 AGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCCCGAGACGTGCCAC 154

Qy        620 GGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACT 679
                   || |||| ||||  | |    |  |||  ||||| |    ||  |        ||
Db        155 CCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGCAGGCAGCT 214
```

```
Qy         680 TTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCC 739
               |||||  | || |||| |||  |  ||||  ||||||   ||||   |||  |  | | |
Db         215 TTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCACAGAACCAGATCGCCAGCCTGC 274

Qy         740 AGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGC 799
                   |     ||| | |   ||||  | |||| ||  |||| | ||      |||  |||
Db         275 CCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGC 334

Qy         800 TGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGA 859
               |||      |     | || ||| |||| ||||||     |||  ||| |||   | ||
Db         335 TGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGG 394

Qy         860 ACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGG 919
                ||| ||||   |  ||  ||||     ||  || || ||     ||   |      |||
Db         395 GCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGG 454

Qy         920 AGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCA 979
               ||   ||  | ||     |||| | || |  |  |||||   |     | |||  |||
Db         455 AG---CTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGCCTGCCCCGCC 511

Qy         980 TGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCA 1039
               || |    ||  |||       |     || |      ||||| | |  |  |  || ||
Db         512 TGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACA 571

Qy        1040 C 1040
               |
Db         572 C 572

RESULT 6
US-09-182-024A-1
; Sequence 1, Application US/09182024A
; Patent No. 6342370
; GENERAL INFORMATION:
;  APPLICANT:  Connolly, Timothy
;  APPLICANT:  Rajput, Bhanu
;  TITLE OF INVENTION:  Human Slit Polypeptide and Polynucleotides Encoding
;  TITLE OF INVENTION:  Same
;  FILE REFERENCE:  640100-271
;  CURRENT APPLICATION NUMBER: US/09/182,024A
;  CURRENT FILING DATE:  1998-10-29
;  PRIOR APPLICATION NUMBER:  60/063,946
;  PRIOR FILING DATE:  1997-10-31
;  PRIOR APPLICATION NUMBER:  60/096,420
;  PRIOR FILING DATE:  1998-08-13
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 5176
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:cDNA coding for
;   OTHER INFORMATION: the human slit polypeptide of SEQ ID NO:2
US-09-182-024A-1

  Query Match              2.5%;  Score 77.2;  DB 4;  Length 5176;
  Best Local Similarity   51.8%;  Pred. No. 9.5e-08;
  Matches  200;  Conservative    0;  Mismatches  183;  Indels    3;  Gaps    1;

Qy         700 GTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGAC 759
               | |  |||   ||||  ||| || |  |||| |   |||        |   |||   |
Db         409 GAGCGCCTTGACCTGGACAGAAATAATATCACCAGGATCACCAAGATGGACTTCGCTGGG 468

Qy         760 CTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGAC 819
               ||| ||| ||||||    | |||||||  ||    ||| | | |        | | | | |
Db         469 CTCAAGAACCTCCGAGTCTTGCATCTGGAAGACAACCAGGTCAGCGTCATCGAGAGAGGC 528

Qy         820 ACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGC 879
                  || ||| || ||||    | ||  |||  ||    | |||||||| ||    |||||   | |
Db         529 GCCTTCCAGGACCTGAAGCAGCTAGAGCGACTGCGCCTGAACAAGAATAAGCTGCAAGTC 588

Qy         880 ATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTAC 939
```

```
                | |  ||||   |  |||  ||     |  || ||   |       || ||  |     |
Db         589 CTTCCAGAATTGCTTTTCCAG---AGCACGCCGAAGCTCACCAGACTAGATTTGAGTGAA 645

Qy         940 AACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTG 999
               ||| |  ||||| ||  | ||| || |  |  | ||  ||||   | |  |  |  | |||
Db         646 AACCAGATCCAGGGGATCCCGAGGAAGGCGTTCCGCGGCATCACCGATGTGAAGAACCTG 705

Qy        1000 AGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAA 1059
                  ||||||||||||   |      ||| |  || ||  ||||   ||  ||||   |
Db         706 CAACTGGACAACAACCACATCAGCTGCATTGAAGATGGAGCCTTCCGAGCGCTGCGCGAT 765

Qy        1060 CTGGCCCGCCTGGATCTCACCTCCAA 1085
                |||    |||    |||| |  |||
Db         766 TTGGAGATCCTTACCCTCAACAACAA 791

RESULT 7
US-08-986-485-1
; Sequence 1, Application US/08986485
; Patent No. 6046030
;  GENERAL INFORMATION:
;    APPLICANT:  WU, SHUJIAN
;    APPLICANT:  SWEET, RAYMOND
;    APPLICANT:  TRUNEH, ALEMSEGED
;    TITLE OF INVENTION:  A HUMAN LIG-1 HOMOLOG (HLIG-1)
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/986,485
;      FILING DATE:  08-DEC-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/059,448
;      FILING DATE:  22-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-70264
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4843 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-986-485-1

  Query Match              2.0%;  Score 63.8;  DB 3;  Length 4843;
  Best Local Similarity   49.1%;  Pred. No. 7.7e-05;
  Matches  198;  Conservative    0;  Mismatches  202;  Indels    3;  Gaps    1;

Qy         684 CAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCC 743
               |||  |  |  |||||    | |||  |||     ||||||  ||  | |  ||  ||
Db         669 CAAGCTACCCAGGCTGACACAACTGGACCTCAATCGGAACAGGATTCGGCTGATAGAGGG 728

Qy         744 CTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCC 803
               | |  |||| | ||  ||| | ||| |       ||| | ||| |  | ||     |
Db         729 CCTCACCTTCCAGGGGCTCAACAGCTTGGAGGTGCTGAAGCTTCAGCGAAACAACATCAG 788
```

```
Qy         804 AAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAA 863
                |  ||    || |   || || |||| |||  |||  ||||    ||   | || |  |
Db         789 CAAACTGACAGATGGGGCCTTCTGGGGACTGTCCAAAATGCATGTGCTGCACCTGGAGTA 848

Qy         864 CAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGA 923
               | ||   ||||  |   |     |   |  |  | | | ||  |    |  || |  |
Db         849 CGACAGCCTGGTAGAAGTGAACAGCGGCTCGCTCTACGGCCTCACGG---CCCTGCATCA 905

Qy         924 TCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGT 983
                ||  |||||  | ||||    | ||  | || | |   |  |   |  | |   |
Db         906 GCTCCACCTCAGCAACAATTCCATCGCTCGCATTCACCGCAAGGGCTGGAGCTTCTGCCA 965

Qy         984 CAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTT 1043
                || || ||  || ||  ||||  |  ||||   |||   |   | | ||||||  | | |
Db         966 GAAGCTGCATGAGTTGGTCCTGTCCTTCAACAACCTGACACGGCTGGACGAGGAGAGCCT 1025

Qy        1044 TGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAAT 1086
                || || |||   |  |||     ||||  ||||| |  ||||
Db        1026 GGCCGAGCTGAGCAGCCTGAGTGTCCTGCGTCTCAGCCACAAT 1068


RESULT 8
US-08-427-497E-4
; Sequence 4, Application US/08427497E
; Patent No. 5969124
;  GENERAL INFORMATION:
;    APPLICANT:  Lemmon, Vance
;    TITLE OF INVENTION:  A Method for Characterizing the
;    TITLE OF INVENTION:  Nucleotide Sequence of L1CAM and
; Patent No. 5969124
;    TITLE OF INVENTION:  the Nucleotide Sequence
;    TITLE OF INVENTION:  Characterized Thereby
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fay, Sharpe, Beall, Fagan,
;      ADDRESSEE:  Minnich & McKee
;      STREET:  1100 Superior Avenue
;      STREET:  Suite 700
;      CITY:  Cleveland
;      STATE:  Ohio
;      COUNTRY:  U.S.A.
;      ZIP:  44114-2518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      MEDIUM TYPE:  storable
;      COMPUTER:  Compaq Prolinea 5100e
;      OPERATING SYSTEM:  DOS 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/427,497E
;      FILING DATE:  April 24, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/904,991
;      FILING DATE:  June 26, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Minnich, Richard J.
;      REGISTRATION NUMBER:  24,175
;      REFERENCE/DOCKET NUMBER:  CWR 2 149-3-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (216) 861-5582
;      TELEFAX:  (216) 241-1666
;      TELEX:  (216) 980162
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2600
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  nucleic acids
;    HYPOTHETICAL:  irrelevant
;    ANTI-SENSE:  no
```

```
;    ORIGINAL SOURCE:
;      ORGANISM:  homo sapiens
;      INDIVIDUAL ISOLATE:  17-18 week fetus
;    IMMEDIATE SOURCE:
;      LIBRARY:  Stratagene cDNA Library 936206
;      CLONE:  4
;    PUBLICATION INFORMATION:
;      AUTHORS:  Hlavin, Mary Louise
;      AUTHORS:  Lemmon, Vance
;      TITLE:  Molecular structure and functional testing of
;      TITLE:  human L1CAM: an interspecies comparison.
;      JOURNAL:  GENOMICS
;      VOLUME:  11
;      ISSUE:
;      PAGES:  416-423
;      DATE:  1991
;      RELEVANT RESIDUES IN SEQ ID NO:  1108 to 3708
US-08-427-497E-4

  Query Match              1.8%;  Score 56.4;  DB 2;  Length 2600;
  Best Local Similarity   49.3%;  Pred. No. 0.0026;
  Matches  147;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy        1334 TCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGT 1393
               || | ||   | |||  || | ||   ||| | |  |||||    | |||   ||   ||
Db         176 TCCTGACTGCGGACAATCAGACGTACATGGCTGTCCAGGGCAGCACTGCCTACCTTCTGT 235

Qy        1394 GCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGG 1453
               |||| ||| | || |  ||    |||  | | || ||| | | |   ||||        |
Db         236 GCAAGGCCTTCGGAGCGCCTGTGCCCAGTGTTCAGTGGCTGGACGAGGATGGGACAACAG 295

Qy        1454 TAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCA 1513
               |     |   |  | |   |     ||||| |||||| ||||||| |||      |  |
Db         296 TGCTTCAGGACGAACGCTTCTTCCCCTATGCCAATGGGACCCTGGGCATTCGAGACCTCC 355

Qy        1514 CATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGG 1573
                  |  | ||||  ||   || |  |||| | ||||||||||         | |  ||
Db         356 AGGCCAATGACACCGGACGCTACTTCTGCCTGGCTGCCAATGACCAAAACAATGTTACCA 415

Qy        1574 CCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCAC 1631
                |||||   |  |     |     |     | ||  |||     |  |  || || ||||
Db         416 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCGCAGCAC 473


RESULT 9
US-08-427-497E-3
; Sequence 3, Application US/08427497E
; Patent No. 5969124
;  GENERAL INFORMATION:
;    APPLICANT:  Lemmon, Vance
;    TITLE OF INVENTION:  A Method for Characterizing the
;    TITLE OF INVENTION:  Nucleotide Sequence of L1CAM and
; Patent No. 5969124
;    TITLE OF INVENTION:  the Nucleotide Sequence
;    TITLE OF INVENTION:  Characterized Thereby
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fay, Sharpe, Beall, Fagan,
;      ADDRESSEE:  Minnich & McKee
;      STREET:  1100 Superior Avenue
;      STREET:  Suite 700
;      CITY:  Cleveland
;      STATE:  Ohio
;      COUNTRY:  U.S.A.
;      ZIP:  44114-2518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      MEDIUM TYPE:  storable
;      COMPUTER:  Compaq Prolinea 5100e
;      OPERATING SYSTEM:  DOS 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/427,497E
```

```
;       FILING DATE:  April 24, 1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/904,991
;       FILING DATE:  June 26, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Minnich, Richard J.
;       REGISTRATION NUMBER:  24,175
;       REFERENCE/DOCKET NUMBER:  CWR 2 149-3-1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (216) 861-5582
;       TELEFAX:  (216) 241-1666
;       TELEX:  (216) 980162
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3189
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  nucleic acids
;     HYPOTHETICAL:  irrelevant
;     ANTI-SENSE:  no
;     ORIGINAL SOURCE:
;       ORGANISM:  homo sapiens
;       INDIVIDUAL ISOLATE:  17-18 week fetus
;     IMMEDIATE SOURCE:
;       LIBRARY:  Stratagene cDNA Library 936206
;       CLONE:  3.1
;     PUBLICATION INFORMATION:
;       AUTHORS:  Hlavin, Mary Louise
;       AUTHORS:  Lemmon, Vance
;       TITLE:  Molecular structure and functional testing of
;       TITLE:  human L1CAM: an interspecies comparison.
;       JOURNAL:  GENOMICS
;       VOLUME:  11
;       ISSUE:
;       PAGES:  416-423
;       DATE:  1991
;       RELEVANT RESIDUES IN SEQ ID NO:  548 to 3736
US-08-427-497E-3

  Query Match             1.8%;  Score 56.4;  DB 2;  Length 3189;
  Best Local Similarity   49.3%;  Pred. No. 0.0028;
  Matches  147;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy       1334 TCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGT 1393
              || | ||   | |||  || | ||   ||| | |  |||||    | |||   ||   ||
Db        736 TCCTGACTGCGGACAATCAGACGTACATGGCTGTCCAGGGCAGCACTGCCTACCTTCTGT 795

Qy       1394 GCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGG 1453
              |||| ||| | || |  ||    |||  | | || ||| | | |   ||||        |
Db        796 GCAAGGCCTTCGGAGCGCCTGTGCCCAGTGTTCAGTGGCTGGACGAGGATGGGACAACAG 855

Qy       1454 TAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCA 1513
              |     |   |  | |   |     ||||| |||||| ||||||| |||      |  |
Db        856 TGCTTCAGGACGAACGCTTCTTCCCCTATGCCAATGGGACCCTGGGCATTCGAGACCTCC 915

Qy       1514 CATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGG 1573
                 |  | ||||  ||   || |  |||| | ||||||||||         | |  ||
Db        916 AGGCCAATGACACCGGACGCTACTTCTGCCTGGCTGCCAATGACCAAAACAATGTTACCA 975

Qy       1574 CCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCAC 1631
               |||||   |  |     |     |     | ||  |||   |  |  || || ||||
Db        976 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCGCAGCAC 1033


RESULT 10
US-08-341-843B-1
; Sequence 1, Application US/08341843B
; Patent No. 5872225
;  GENERAL INFORMATION:
;     APPLICANT:  Lemmon, Vance
;     TITLE OF INVENTION:  A Method for Characterizing the
;     TITLE OF INVENTION:  Nucleotide Sequence of L1CAM and
; Patent No. 5872225
;     TITLE OF INVENTION:  the Nucleotide Sequence
;     TITLE OF INVENTION:  Characterized Thereby
;     NUMBER OF SEQUENCES:  39
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fay, Sharpe, Beall, Fagan,
;       ADDRESSEE:  Minnich & McKee
;       STREET:  1100 Superior Avenue
;       STREET:  Suite 700
;       CITY:  Cleveland
;       STATE:  Ohio
;       COUNTRY:  U.S.A.
;       ZIP:  44114-2518
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;       MEDIUM TYPE:  storable
;       COMPUTER:  Compaq Prolinea 5100e
;       OPERATING SYSTEM:  DOS 5.0
;       SOFTWARE:  ASCII
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/341,843B
;       FILING DATE:  No. 5872225ember 18, 1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/904,991
;       FILING DATE:  June 26, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Minnich, Richard J.
;       REGISTRATION NUMBER:  24,175
;       REFERENCE/DOCKET NUMBER:  CWR 2 149-1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (216) 861-5582
;       TELEFAX:  (216) 241-1666
;       TELEX:  (216) 980162
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3774
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  irrelevant
;     ANTI-SENSE:  no
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo Sapiens
;       INDIVIDUAL ISOLATE:  17-18 week fetus
;     IMMEDIATE SOURCE:
;       LIBRARY:  Stratagene cDNA Library 936206
;       CLONE:  synthesis of 4 clones
;     PUBLICATION INFORMATION:
;       AUTHORS:  Hlavin, Mary Louise
;       AUTHORS:  Lemmon, Vance
;       TITLE:  Molecular structure and functional
;       TITLE:  testing of human L1CAM: an
;       TITLE:  interspecies comparison.
;       JOURNAL:  GENOMICS
;       VOLUME:  11
;       ISSUE:
;       PAGES:  416-423
;       DATE:  1991
;       RELEVANT RESIDUES IN SEQ ID NO:  1 to 3774
US-08-341-843B-1

  Query Match             1.8%;  Score 56.4;  DB 2;  Length 3774;
  Best Local Similarity   49.3%;  Pred. No. 0.0029;
  Matches  147;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy       1334 TCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGT 1393
              || | ||   | |||  || | ||   ||| | |  |||||    | |||   ||   ||
Db       1283 TCCTGACTGCGGACAATCAGACGTACATGGCTGTCCAGGGCAGCACTGCCTACCTTCTGT 1342

Qy       1394 GCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGG 1453
```

```
              |||| ||| |  || |  || |   |||  | || ||| |  | |||  |       |
Db       1343 GCAAGGCCTTCGGAGCGCCTGTGCCCAGTGTTCAGTGGCTGGACGAGGATGGGACAACAG 1402

Qy       1454 TAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCA 1513
              |     |   |  | |   |     ||||| |||||| |||||||| |||      |  |
Db       1403 TGCTTCAGGACGAACGCTTCTTCCCCTATGCCAATGGGACCCTGGGCATTCGAGACCTCC 1462

Qy       1514 CATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGG 1573
                 |  | ||||  ||   || |  |||| | ||||||||||         | |  ||
Db       1463 AGGCCAATGACACCGGACGCTACTTCTGCCTGGCTGCCAATGACCAAAACAATGTTACCA 1522

Qy       1574 CCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCAC 1631
               |||||   |  |     |     |     | ||  |||   |  |  || ||  ||||
Db       1523 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCGCAGCAC 1580


RESULT 11
US-08-427-497E-1
; Sequence 1, Application US/08427497E
; Patent No. 5969124
;  GENERAL INFORMATION:
;    APPLICANT:  Lemmon, Vance
;    TITLE OF INVENTION:  A Method for Characterizing the
;    TITLE OF INVENTION:  Nucleotide Sequence of L1CAM and
; Patent No. 5969124
;    TITLE OF INVENTION:  the Nucleotide Sequence
;    TITLE OF INVENTION:  Characterized Thereby
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fay, Sharpe, Beall, Fagan,
;      ADDRESSEE:  Minnich & McKee
;      STREET:  1100 Superior Avenue
;      STREET:  Suite 700
;      CITY:  Cleveland
;      STATE:  Ohio
;      COUNTRY:  U.S.A.
;      ZIP:  44114-2518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      MEDIUM TYPE:  storable
;      COMPUTER:  Compaq Prolinea 5100e
;      OPERATING SYSTEM:  DOS 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/427,497E
;      FILING DATE:  April 24, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/904,991
;      FILING DATE:  June 26, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Minnich, Richard J.
;      REGISTRATION NUMBER:  24,175
;      REFERENCE/DOCKET NUMBER:  CWR 2 149-3-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (216) 861-5582
;      TELEFAX:  (216) 241-1666
;      TELEX:  (216) 980162
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3774
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  nucleic acids
;    HYPOTHETICAL:  irrelevant
;    ANTI-SENSE:  no
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo Sapiens
;      INDIVIDUAL ISOLATE:  17-18 week fetus
;    IMMEDIATE SOURCE:
;      LIBRARY:  Stratagene cDNA Library 936206
;      CLONE:  synthesis of 4 clones
;    PUBLICATION INFORMATION:
;      AUTHORS:  Hlavin, Mary Louise
;      AUTHORS:  Lemmon, Vance
;      TITLE:  Molecular structure and functional
;      TITLE:  testing of human L1CAM: an
;      TITLE:  interspecies comparison.
;      JOURNAL:  GENOMICS
;      VOLUME:  11
;      ISSUE:
;      PAGES:  416-423
;      DATE:  1991
;      RELEVANT RESIDUES IN SEQ ID NO:  1 to 3774
US-08-427-497E-1

  Query Match             1.8%;  Score 56.4;  DB 2;  Length 3774;
  Best Local Similarity   49.3%;  Pred. No. 0.0029;
  Matches  147;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy       1334 TCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGT 1393
              || | ||   | |||  || | ||   ||| | |  |||||    | |||   ||   ||
Db       1283 TCCTGACTGCGGACAATCAGACGTACATGGCTGTCCAGGGCAGCACTGCCTACCTTCTGT 1342

Qy       1394 GCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGG 1453
              |||| ||| | || |  ||    |||  | | || ||| | | |   ||||        |
Db       1343 GCAAGGCCTTCGGAGCGCCTGTGCCCAGTGTTCAGTGGCTGGACGAGGATGGGACAACAG 1402

Qy       1454 TAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCA 1513
              |     |   |  | |   |     ||||| |||||| |||||||| |||      |  |
Db       1403 TGCTTCAGGACGAACGCTTCTTCCCCTATGCCAATGGGACCCTGGGCATTCGAGACCTCC 1462

Qy       1514 CATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGG 1573
                 |  | ||||  ||   || |  |||| | ||||||||||         | |  ||
Db       1463 AGGCCAATGACACCGGACGCTACTTCTGCCTGGCTGCCAATGACCAAAACAATGTTACCA 1522

Qy       1574 CCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCAC 1631
               |||||   |  |     |     |     | ||  |||   |  |  || ||  ||||
Db       1523 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCGCAGCAC 1580


RESULT 12
US-08-427-497E-2
; Sequence 2, Application US/08427497E
; Patent No. 5969124
;  GENERAL INFORMATION:
;    APPLICANT:  Lemmon, Vance
;    TITLE OF INVENTION:  A Method for Characterizing the
;    TITLE OF INVENTION:  Nucleotide Sequence of L1CAM and
; Patent No. 5969124
;    TITLE OF INVENTION:  the Nucleotide Sequence
;    TITLE OF INVENTION:  Characterized Thereby
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fay, Sharpe, Beall, Fagan,
;      ADDRESSEE:  Minnich & McKee
;      STREET:  1100 Superior Avenue
;      STREET:  Suite 700
;      CITY:  Cleveland
;      STATE:  Ohio
;      COUNTRY:  U.S.A.
;      ZIP:  44114-2518
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      MEDIUM TYPE:  storable
;      COMPUTER:  Compaq Prolinea 5100e
;      OPERATING SYSTEM:  DOS 5.0
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/427,497E
;      FILING DATE:  April 24, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/904,991
;      FILING DATE:  June 26, 1992
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Minnich, Richard J.
;      REGISTRATION NUMBER:  24,175
;      REFERENCE/DOCKET NUMBER:  CWR 2 149-3-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (216) 861-5582
;      TELEFAX:  (216) 241-1666
;      TELEX:  (216) 980162
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3774
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  nucleic acids
;    HYPOTHETICAL:  irrelevant
;    ANTI-SENSE:  no
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo Sapiens
;      INDIVIDUAL ISOLATE:  17-18 week fetus
;    IMMEDIATE SOURCE:
;      LIBRARY:  Stratagene cDNA Library 936206
;      CLONE:  synthesis of 4 clones
;    PUBLICATION INFORMATION:
;      AUTHORS:  Hlavin, Mary Louise
;      AUTHORS:  Lemmon, Vance
;      TITLE:  Molecular structure and functional testing of
;      TITLE:  human L1CAM: an interspecies comparison.
;      JOURNAL:  GENOMICS
;      VOLUME:  11
;      ISSUE:
;      PAGES:  416-423
;      DATE:  1991
;      RELEVANT RESIDUES IN SEQ ID NO:  1 to 3774
US-08-427-497E-2

  Query Match               1.8%;  Score 56.4;  DB 2;  Length 3774;
  Best Local Similarity   49.3%;  Pred. No. 0.0029;
  Matches  147;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy       1334 TCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGT 1393
              || | ||   | |||  || | ||   ||| | |  |||||    | |||   ||   ||
Db       1283 TCCTGACTGCGGACAATCAGACGTACATGGCTGTCCAGGGCAGCACTGCCTACCTTCTGT 1342

Qy       1394 GCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGG 1453
              |||| ||| | || |  ||    |||  | | || ||| | | |   ||||        |
Db       1343 GCAAGGCCTTCGGAGCGCCTGTGCCCAGTGTTCAGTGGCTGGACGAGGATGGGACAACAG 1402

Qy       1454 TAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCA 1513
              |     |   |  | |   |     ||||| |||||| ||||||| |||      |  |
Db       1403 TGCTTCAGGACGAACGCTTCTTCCCCTATGCCAATGGGACCCTGGGCATTCGAGACCTCC 1462

Qy       1514 CATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGG 1573
                 |  | ||||  ||   || |  |||| | ||||||||||         | |  ||
Db       1463 AGGCCAATGACACCGGACGCTACTTCTGCCTGGCTGCCAATGACCAAAACAATGTTACCA 1522

Qy       1574 CCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCAC 1631
               |||||   |  |     |     |     | ||  |||   |  |  || || ||||
Db       1523 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCGCAGCAC 1580


RESULT 13
US-08-506-296B-13
; Sequence 13, Application US/08506296B
; Patent No. 6313265
;  GENERAL INFORMATION:
;    APPLICANT:  Phillips, Greg
;    APPLICANT:  Cunningham, Bruce A.
;    APPLICANT:  Crossin, Kathryn L.
;    TITLE OF INVENTION:  NEURITE OUTGROWTH-PROMOTING POLYPEPTIDES
;    TITLE OF INVENTION:  CONTAINING FIBRONECTIN TYPE III REPEATS AND METHODS OF USE
;    NUMBER OF SEQUENCES:  77
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute
;      STREET:  10550 No. 6313265th Torrey Pines Road, TPC-8
;      CITY:  La Jolla
;      STATE:  California
;      COUNTRY:  U.S.
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/506,296B
;      FILING DATE:  24-JUL-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI 488.0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 554-2937
;      TELEFAX:  (619) 554-6312
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3888 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  12..3773
US-08-506-296B-13

  Query Match               1.8%;  Score 56.4;  DB 4;  Length 3888;
  Best Local Similarity   49.3%;  Pred. No. 0.0029;
  Matches  147;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy       1334 TCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGT 1393
              || | ||   | |||  || | ||   ||| | |  |||||    | |||   ||   ||
Db       1294 TCCTGACTGCGGACAATCAGACGTACATGGCTGTCCAGGGCAGCACTGCCTACCTTCTGT 1353

Qy       1394 GCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGG 1453
              |||| ||| | || |  ||    |||  | | || ||| | | |   ||||        |
Db       1354 GCAAGGCCTTCGGAGCGCCTGTGCCCAGTGTTCAGTGGCTGGACGAGGATGGGACAACAG 1413

Qy       1454 TAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCA 1513
              |     |   |  | |   |     ||||| |||||| ||||||| |||      |  |
Db       1414 TGCTTCAGGACGAACGCTTCTTCCCCTATGCCAATGGGACCCTGGGCATTCGAGACCTCC 1473

Qy       1514 CATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGG 1573
                 |  | ||||  ||   || |  |||| | ||||||||||         | |  ||
Db       1474 AGGCCAATGACACCGGACGCTACTTCTGCCTGGCTGCCAATGACCAAAACAATGTTACCA 1533

Qy       1574 CCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCAC 1631
               |||||   |  |     |     |     | ||  |||   |  |  || || ||||
Db       1534 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCGCAGCAC 1591


RESULT 14
US-08-866-757-1
; Sequence 1, Application US/08866757
; Patent No. 5858716
;  GENERAL INFORMATION:
;    APPLICANT:  ELSHOURBAGY, NABIL A
;    APPLICANT:  LI, XIAOTONG
;    APPLICANT:  BERGSMA, DERK J
;    TITLE OF INVENTION:  NOVEL 7TM RECEPTOR (H2CAA71)
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
```

```
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/866,757
;      FILING DATE:  30-MAY-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-70055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4203 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-866-757-1

  Query Match             1.6%;  Score 51.6;  DB 2;  Length 4203;
  Best Local Similarity   54.1%;  Pred. No. 0.033;
  Matches  105;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy        684 CAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCC 743
              |||  || |    ||   ||  ||||||||| |    ||||  |||     |||||
Db         85 CAATCTGCCCACCCTACAGGCGCTGACCCTGGCTCTCAACAAGATCTCAAGCATCCCTGA 144

Qy        744 CTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCC 803
              |||| | |||    ||||    |||||       ||||||||| | | ||||    |
Db        145 CTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGCATCTTCATAACAATAAAATTAG 204

Qy        804 AAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAA 863
              ||||||  |  |  ||    |    || || |  |||||| ||  | |   | ||||  |
Db        205 AAGCCTGAGTCAACACTGTTTTGATGGACTAGATAACCTGGAGACCTTAGACTTGAATTA 264

Qy        864 CAACCAGCTGGGCG 877
               ||  |  |||| |
Db        265 TAATAACTTGGGGG 278

RESULT 15
US-09-153-593-1
; Sequence 1, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
;  APPLICANT: ELSHOURBAGY, NABIL A
;  APPLICANT: LI, XIAOTONG
;  APPLICANT: BERGSMA, DERK J
;  TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
;  FILE REFERENCE: GH-70055-1
;  CURRENT APPLICATION NUMBER: US/09/153,593A
;  CURRENT FILING DATE: 1998-09-15
;  EARLIER APPLICATION NUMBER: 08/866,757
;  EARLIER FILING DATE: 1997-05-30
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4203
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-153-593-1

  Query Match             1.6%;  Score 51.6;  DB 4;  Length 4203;
  Best Local Similarity   54.1%;  Pred. No. 0.033;
  Matches  105;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy        684 CAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCC 743
              |||  || |    ||   ||  ||||||||| |    ||||  |||     |||||
Db         85 CAATCTGCCCACCCTACAGGCGCTGACCCTGGCTCTCAACAAGATCTCAAGCATCCCTGA 144

Qy        744 CTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCC 803
              |||| | |||    ||||    |||||       ||||||||| | | ||||    |
Db        145 CTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGCATCTTCATAACAATAAAATTAG 204

Qy        804 AAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAA 863
              ||||||  |  |  ||    |    || || |  |||||| ||  | |   | ||||  |
Db        205 AAGCCTGAGTCAACACTGTTTTGATGGACTAGATAACCTGGAGACCTTAGACTTGAATTA 264

Qy        864 CAACCAGCTGGGCG 877
               ||  |  |||| |
Db        265 TAATAACTTGGGGG 278

Search completed: June  8, 2003, 19:06:49
Job time : 237 secs
```

```
                     GenCore version 5.1.6
            Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         June  8, 2003, 18:21:58 ; Search time 411 Seconds
                                           (without alignments)
                                           10705.950 Million cell updates/sec

Title:          US-09-831-846-1
Perfect score:  3144
Sequence:       1 gcctggctccctctcgctga..........gcggctttggattgcttatg 3144

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters:      1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                    %
Result              Query
   No.    Score     Match Length DB  ID                        Description
-------------------------------------------------------------------------------
     1      633     20.1    1887  10  US-09-815-626-3            Sequence 3, Appli
     2      633     20.1    2558  10  US-09-815-626-1            Sequence 1, Appli
     3    621.8     19.8    2855  9  US-10-245-103-99           Sequence 99, Appl
     4    621.8     19.8    2855  9  US-10-245-107-99           Sequence 99, Appl
     5    621.8     19.8    2855  9  US-10-245-143-99           Sequence 99, Appl
     6    621.8     19.8    2855  9  US-10-245-771-99           Sequence 99, Appl
     7    621.8     19.8    2855  9  US-10-245-851-99           Sequence 99, Appl
     8    621.8     19.8    2855  9  US-10-245-883-99           Sequence 99, Appl
     9    621.8     19.8    2855  9  US-10-237-535-99           Sequence 99, Appl
    10    621.8     19.8    2855  9  US-10-238-183-99           Sequence 99, Appl
    11    621.8     19.8    2855  9  US-10-238-283-99           Sequence 99, Appl
    12    621.8     19.8    2855  9  US-10-238-370-99           Sequence 99, Appl
    13    621.8     19.8    2855  9  US-10-245-055-99           Sequence 99, Appl
    14    621.8     19.8    2855  9  US-10-245-147-99           Sequence 99, Appl
    15    621.8     19.8    2855  9  US-10-245-730-99           Sequence 99, Appl
    16    621.8     19.8    2855  9  US-10-245-739-99           Sequence 99, Appl
    17    621.8     19.8    2855  9  US-10-246-210-99           Sequence 99, Appl
    18    621.8     19.8    2855  9  US-10-239-196-99           Sequence 99, Appl
    19    621.8     19.8    2855  9  US-10-243-024-99           Sequence 99, Appl
    20    621.8     19.8    2855  9  US-10-243-409-99           Sequence 99, Appl
    21    621.8     19.8    2855  9  US-10-245-033-99           Sequence 99, Appl
    22    621.8     19.8    2855  9  US-10-245-621-99           Sequence 99, Appl
    23    621.8     19.8    2855  9  US-10-245-880-99           Sequence 99, Appl
    24    621.8     19.8    2855  9  US-10-243-095-99           Sequence 99, Appl
    25    621.8     19.8    2855  9  US-10-245-185-99           Sequence 99, Appl
    26    621.8     19.8    2855  9  US-10-245-427-99           Sequence 99, Appl
    27    621.8     19.8    2855  9  US-10-245-473-99           Sequence 99, Appl
    28    621.8     19.8    2855  9  US-10-245-770-99           Sequence 99, Appl
    29    621.8     19.8    2855  9  US-10-245-877-99           Sequence 99, Appl
    30    621.8     19.8    2855  9  US-10-246-976-99           Sequence 99, Appl
    31    621.8     19.8    2855  9  US-10-243-320-99           Sequence 99, Appl
    32    621.8     19.8    2855  9  US-10-242-743-99           Sequence 99, Appl
    33    621.8     19.8    2855  9  US-10-242-845-99           Sequence 99, Appl
    34      608     19.3    1653  10  US-09-897-214-7             Sequence 7, Appli
    35    554.6     17.6    1653  10  US-09-897-214-9             Sequence 9, Appli
    36    477.2     15.2    8269  9  US-10-125-540-551          Sequence 551, App
    37    477.2     15.2    8269  9  US-10-073-961-452          Sequence 452, App
    38    477.2     15.2    8269  9  US-10-103-313-597          Sequence 597, App
    39    477.2     15.2    8269  10  US-09-764-887-452          Sequence 452, App
    40    477.2     15.2    8269  10  US-09-764-870-551          Sequence 551, App
    41    477.2     15.2    8269  10  US-09-764-853-879          Sequence 879, App
    42    461.4     14.7    1872  10  US-09-768-826-21           Sequence 21, Appl
    43    273.6      8.7     493  9  US-09-918-995-26236        Sequence 26236, A
    44    204.2      6.5     432  10  US-09-897-214-4             Sequence 4, Appli
    45    204.2      6.5     432  10  US-09-893-737-189          Sequence 189, App

                              ALIGNMENTS


RESULT 1
US-09-815-626-3
; Sequence 3, Application US/09815626
; Patent No. US20020076752A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria Alexandra
;  TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT
;  TITLE OF INVENTION:  FAMILY MEMBER AND USES THEREOF
;  FILE REFERENCE: 10448-031001
;  CURRENT APPLICATION NUMBER: US/09/815,626
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: US 60/191,863
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 1887
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-815-626-3

  Query Match             20.1%;  Score 633;  DB 10;  Length 1887;
  Best Local Similarity   61.3%;  Pred. No. 9e-166;
  Matches 1078;  Conservative    0;  Mismatches  665;  Indels   16;  Gaps    3;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
              ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db         63 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 122

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                |    ||||||||       || || || || || || ||     | ||||| ||| |
Db        123 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 182

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
              ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db        183 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 242

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
              |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db        243 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 302

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
```

```
Db        303 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 362

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866

Db        363 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 422

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926

Db        423 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 482

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986

Db        483 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 542

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046

Db        543 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 602

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106

Db        603 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 662

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166

Db        663 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 722

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226

Db        723 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 782

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286

Db        783 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 842

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346

Db        843 CTACTTTTGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 902

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406

Db        903 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 962

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466

Db        963 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1022

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526

Db       1023 CCGTGCCCGCGCCTTCCCCAATGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1082

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586

Db       1083 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1142

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644

Db       1143 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1202

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704

Db       1203 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1262

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764

Db       1263 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1314

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824

Db       1315 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1374

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884

Db       1375 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1434

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944

Db       1435 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1494

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004

Db       1495 GTGTATGAGGGCAGTGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1554

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064

Db       1555 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 1614

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124

Db       1615 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 1674

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178

Db       1675 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCC 1734

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238

Db       1735 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 1794

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257

Db       1795 CCCGCCCCGGAGCCCGCGG 1813


RESULT 2
US-09-815-626-1
; Sequence 1, Application US/09815626
; Patent No. US20020076752A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria Alexandra
;  TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT
;  TITLE OF INVENTION:  FAMILY MEMBER AND USES THEREOF
;  FILE REFERENCE: 10448-031001
;  CURRENT APPLICATION NUMBER: US/09/815,626
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: US 60/191,863
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2558
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (435)...(2318)
US-09-815-626-1

  Query Match             20.1%;  Score 633;  DB 10;  Length 2558;
  Best Local Similarity   61.3%;  Pred. No. 9.9e-166;
  Matches 1078;  Conservative    0;  Mismatches  665;  Indels   16;  Gaps    3;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566

Db        497 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 556

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626

Db        557 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 616

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686

Db        617 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 676

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
```

```
Db         677 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 736

Qy         747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                 ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db         737 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 796

Qy         807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
                || || |||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db         797 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 856

Qy         867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
               ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db         857 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 916

Qy         927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
                ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db         917 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 976

Qy         987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
               | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db         977 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCCGCCGGCGCTTTTTC 1036

Qy        1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                  |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db        1037 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1096

Qy        1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
                || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db        1097 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1156

Qy        1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
               | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db        1157 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1216

Qy        1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                 | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db        1217 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1276

Qy        1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
               |||||| |||   ||| | |||||||||||||| |||||||||||  |  | || || |
Db        1277 CTACTTTTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1336

Qy        1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
               | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db        1337 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1396

Qy        1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
               ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db        1397 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1456

Qy        1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                 |  ||   | ||    |||||| || |||||  |  |  |||||    |    ||  |
Db        1457 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1516

Qy        1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
               |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db        1517 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1576

Qy        1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                 |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db        1577 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1636

Qy        1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                  ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db        1637 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1696

Qy        1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
                |      | | ||| |        | || || |  || |  ||| ||||        ||
Db        1697 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 1748

Qy        1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
               ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db        1749 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 1808

Qy        1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
                | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db        1809 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 1868

Qy        1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                   | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db        1869 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 1928

Qy        1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
                |||  || | ||  ||||||   |||||||||||       |||||||||||||  |||
Db        1929 GTGTATGAGGGCAGTGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 1988

Qy        2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
               | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db        1989 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2048

Qy        2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
               |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db        2049 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2108

Qy        2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
                |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| ||
Db        2109 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCC 2168

Qy        2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
               || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db        2169 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2228

Qy        2239 GCCGGGGCCCCGCCGCAGG 2257
                |||   |   |||   ||
Db        2229 CCCGCCCCGGAGCCCGCGG 2247

RESULT 3
US-10-245-103-99
; Sequence 99, Application US/10245103
; Publication No. US20030068778A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C112
;  CURRENT APPLICATION NUMBER: US/10/245,103
;  CURRENT FILING DATE:  2002-09-17
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
```

```
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-103-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
Db        798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
Db        858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
Db        918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
Db        978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
Db       1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
Db       1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
Db       1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
Db       1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
Db       1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
Db       1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
Db       1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
Db       1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
Db       1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
Db       1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
Db       1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
Db       1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
Db       1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
Db       1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
Db       1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
Db       1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
Db       1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
Db       2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
Db       2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
Db       2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
Db       2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
Db       2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
Db       2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
Db       2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
Db       2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257
Db       2527 CCCGCCCCGGAGCCCGCGG 2545


RESULT 4
US-10-245-107-99
; Sequence 99, Application US/10245107
; Publication No. US20030068779A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
```

```
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C71
;  CURRENT APPLICATION NUMBER: US/10/245,107
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-107-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566

Db        798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626

Db        858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686

Db        918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746

Db        978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806

Db       1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866

Db       1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926

Db       1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986

Db       1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046

Db       1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106

Db       1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166

Db       1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226

Db       1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286

Db       1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346

Db       1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406

Db       1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466

Db       1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526

Db       1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586

Db       1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644

Db       1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704

Db       1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764

Db       1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824

Db       2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884

Db       2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944

Db       2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004

Db       2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
```

```
Db   2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy   2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
          |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db   2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy   2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
           |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db   2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy   2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
          || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db   2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy   2239 GCCGGGGCCCCGCCGCAGG 2257
           |||   |   |||   ||
Db   2527 CCCGCCCCGGAGCCCGCGG 2545


RESULT 5
US-10-245-143-99
; Sequence 99, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C90
;  CURRENT APPLICATION NUMBER: US/10/245,143
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-143-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;
```

```
Qy    507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
          ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db    798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy    567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
            |    ||||||||       || || || || || || ||     | ||||| ||| |
Db    858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy    627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
          ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db    918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy    687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
          |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db    978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy    747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
            ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db   1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy    807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
           || || |||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db   1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy    867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
          ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db   1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy    927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
           ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db   1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy    987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
          | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db   1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy   1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
             |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db   1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy   1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
           || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db   1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy   1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
          | ||    |    ||||| || ||||| || |||||||| || |||||  |  |||||||
Db   1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy   1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
            | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db   1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy   1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
          ||||||||||   ||| | |||||||||||||| |||||||||||  |  | || || |
Db   1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy   1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
          | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db   1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy   1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
          ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db   1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy   1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
            |  ||   | ||    |||||| || |||||  |  |  |||||    |    ||  |
Db   1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy   1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
          |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db   1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy   1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
```

```
               |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db      1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy      1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db      1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy      1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
              |      | | ||| |        | || || |  || |  ||| ||||        ||
Db      1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy      1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
             ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db      2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy      1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
              | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db      2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy      1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                 | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db      2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy      1945 ATGTGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
              |||  || ||||  ||||||   |||||||||||       |||||||||||||  |||
Db      2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy      2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
             | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db      2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy      2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
             |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db      2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy      2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
              |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db      2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy      2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
             || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db      2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy      2239 GCCGGGGCCCCGCCGCAGG 2257
              |||   |   |||   ||
Db      2527 CCCGCCCCGGAGCCCGCGG 2545


RESULT 6
US-10-245-771-99
; Sequence 99, Application US/10245771
; Publication No. US20030068781A1
; GENERAL INFORMATION:
;  APPLICANT:  Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C98
;  CURRENT APPLICATION NUMBER: US/10/245,771
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-771-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy       507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
             ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db       798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy       567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
               |    ||||||||       || || || || || ||  ||     | ||||| ||| |
Db       858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy       627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
             ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db       918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy       687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
             |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db       978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy       747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
               ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db      1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy       807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
              || || |||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db      1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy       867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
             ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db      1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy       927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
              ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db      1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy       987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
             | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db      1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy      1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db      1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy      1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
              || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db      1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454
```

```
Qy      1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
             | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db      1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy      1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
               | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db      1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy      1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
             ||||||||||   |||   |||||||||||||||| ||||||||||  |  | || || | 
Db      1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy      1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
             | |||      ||      |  |  || | | ||||  | ||    |||   ||  | ||
Db      1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy      1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
             ||||||    |||| | | |  |||||  | ||| | |  || |   ||||| |||||   
Db      1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy      1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
               |  ||   | ||    |||||| || ||||||  |  |  |||||    |    ||  |
Db      1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy      1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
             |||   |||||||||||||| |||||||| || |  |||||||| ||    |||||| |
Db      1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy      1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
               |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db      1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy      1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                ||    ||      | ||  | ||| |     | |||   |     |     |||| 
Db      1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy      1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
              |      | | ||| |        | || || |  || |  ||| ||||        ||
Db      1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy      1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
             ||  | || || ||| |||||| | |       | |    ||  |  |  |    |||||||||
Db      2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy      1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
              | ||||||||| |||     |||||  | ||  | |||||||||||||| ||    | |
Db      2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy      1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                 | ||| || | |   |||||| ||||||      ||||||  ||||  ||||| |||
Db      2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy      1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
              |||  || ||||  ||||||   |||||||||||       |||||||||||||  |||
Db      2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy      2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
             | |||| |  |||     | ||  |||  | |   ||||   |   | ||||||||||| 
Db      2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy      2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
             |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | || 
Db      2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy      2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
              |  |||||||||||||   |  |  |  |||| ||||| |      ||||||  | ||| | 
Db      2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy      2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
             || ||||  || | ||| ||||||||||||||||| |   ||  | || |    ||| || 
Db      2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy      2239 GCCGGGGCCCCGCCGCAGG 2257
              |||   |   |||   ||
Db      2527 CCCGCCCCGGAGCCCGCGG 2545

RESULT 7
US-10-245-851-99
; Sequence 99, Application US/10245851
; Publication No. US20030068782A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C93
;  CURRENT APPLICATION NUMBER: US/10/245,851
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-851-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy       507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
             ||  |||     | |  | || |||     |||   ||||||||  | ||   | | ||
Db       798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy       567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
               |    ||||||||       || |  || || || || ||     | |||| ||| |
Db       858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy       627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
             ||  |||||||| |||   | ||||||||||  |  | |  ||||||  ||| | |||||
Db       918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy       687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
             |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||  
Db       978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037
```

```
Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db       1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
               || || ||||| |   || || ||||||||||||| |||  ||||| ||| | | ||||||
Db       1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
              ||||||||  |   | || |  |  ||  |  || || |       |||| ||||||| ||
Db       1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
               ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db       1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
              | || ||  | || |||| |||||||||||||||| |   |    |  ||   ||| | ||| |
Db       1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                 |||||| || ||||||||| |||||  | ||||||||| || |||   |   | || ||
Db       1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
               || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db       1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
              | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db       1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                | || | ||||||| || |||| ||  ||||| |  ||||||   |  ||    || ||
Db       1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
              ||||||||||   ||| | |||||||||||||||| ||||||||||||||  |  | || || |
Db       1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
              | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db       1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
              ||||||    |||| | | |  |||||  | ||||   | || ||| ||||| |||||
Db       1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                |  ||   | ||    |||||| || |||||  |  |   |||||    |    ||  |
Db       1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
              |||   |||||||||||||| |||||||||| || || |||||||| ||    |||||| |
Db       1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db       1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                 ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db       1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
               |      | | ||| |        | || || |  || |  ||| ||||        ||
Db       1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
              ||  | || || ||| |||||| | |       | |    ||  |  |    |||||||||
Db       2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
               | ||||||||| |||| | |||||  | ||  | |||||||||||||| ||    | |
Db       2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                  | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db       2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
               |||  || ||||  ||||||   |||||||||||       |||||||||||||  |||
Db       2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
              | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db       2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
              |||||| | |   | || ||| ||||||| ||| || | ||||||||||||| || | ||
Db       2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
               |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db       2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
              || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db       2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257
               |||   |   |||   ||
Db       2527 CCCGCCCCGGAGCCCGCGG 2545

RESULT 8
US-10-245-883-99
; Sequence 99, Application US/10245883
; Publication No. US20030068783A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION:  SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C70
;  CURRENT APPLICATION NUMBER: US/10/245,883
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
```

```
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-883-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy       507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
             ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db       798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy       567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
               |    ||||||||       || || || || || || ||     | |||| ||| |
Db       858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy       627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
             ||  |||||||| ||||  | |||||||||||  |  | |  ||||||  ||| | |||||
Db       918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy       687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
             |||||| || ||| || | |||| ||||||  ||||||||||||| ||||| |     ||  
Db       978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy       747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
               ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |   
Db      1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy       807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
              || || ||| |   || || ||||||||||||| |||  ||||| ||| | | ||||||
Db      1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy       867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
             ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db      1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy       927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
              ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db      1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy       987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
             | || ||  | || |||| |||||||||||||||| |     |  ||   ||| | ||| |
Db      1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy      1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                ||||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db      1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy      1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
              || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db      1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy      1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
             | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db      1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy      1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
               | || | |||||| || ||||| ||  ||||| |  ||||||   |  ||    || ||
Db      1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy      1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
             ||||||||||   ||| | |||||||||||||| ||||||||||||  |  | || || | 
Db      1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy      1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
             | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db      1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy      1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
             ||||||    |||| | | |  |||||  | ||| | | ||| ||| ||||| |||||    
Db      1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy      1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
               |  ||   | ||    ||||||| || ||||  |  |  |||||    |    ||  |
Db      1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy      1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
             |||   |||||||||||||| |||||||| || || ||||||| ||    |||||| |
Db      1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy      1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
               |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db      1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy      1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db      1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy      1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
              |      | | ||| |        | || || |  || |  ||| ||||        ||
Db      1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy      1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
             ||  | || || ||| |||||| | |       | |    ||  |  |   |  |||||||||
Db      2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy      1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
              | ||||||||| |||| |  |||||  | ||  | ||||||||||||||| ||    | |
Db      2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy      1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                 | ||| || | |   ||||| ||||||      ||||||  |||| ||||| |||
Db      2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy      1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
              |||  || ||||  |||||||   |||||||||||||       ||||||||||||||  |||
Db      2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy      2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
             | |||| |  |||     | ||  |||  | |   ||||   |   | ||||||||||| 
Db      2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy      2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
             |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | || 
Db      2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy      2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
              |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| | 
Db      2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy      2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
             || ||||  || | ||| ||||||||||||||||| |   ||  | || | |    ||| || 
Db      2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy      2239 GCCGGGGCCCCGCCGCAGG 2257
              |||   |   |||   ||
Db      2527 CCCGCCCCGGAGCCCGCGG 2545

RESULT 9
US-10-237-535-99
; Sequence 99, Application US/10237535
; Publication No. US20030073188A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
```

```
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C3
;  CURRENT APPLICATION NUMBER: US/10/237,535
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/091358
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091978
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/099803
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/106932
;  PRIOR FILING DATE: 1998-11-03
;  PRIOR APPLICATION NUMBER: 60/115554
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/119342
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: 60/123957
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123972
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: 60/127372
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/131271
;  PRIOR FILING DATE: 1999-04-27
;  PRIOR APPLICATION NUMBER: 60/133459
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/135725
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/135729
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/135750
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/138385
;  PRIOR FILING DATE: 1999-06-09
;  PRIOR APPLICATION NUMBER: 60/140653
;  PRIOR FILING DATE: 1999-06-22
;  PRIOR APPLICATION NUMBER: 60/141037
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: 60/144732
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/144758
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/144790
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/145228
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/145698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: 60/146222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: 60/146843
;  PRIOR FILING DATE: 1999-08-03
;  PRIOR APPLICATION NUMBER: 60/148188
;  PRIOR FILING DATE: 1999-08-10
;  PRIOR APPLICATION NUMBER: 60/148513
;  PRIOR FILING DATE: 1999-08-12
;  PRIOR APPLICATION NUMBER: 60/149327
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/149395
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/150114
;  PRIOR FILING DATE: 1999-08-20
;  PRIOR APPLICATION NUMBER: 60/151700
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: 60/151734
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: 60/162506
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: 60/170262
;  PRIOR FILING DATE: 1999-12-09
;  PRIOR APPLICATION NUMBER: 60/177118
;  PRIOR FILING DATE: 2000-01-20
;  PRIOR APPLICATION NUMBER: 60/179851
;  PRIOR FILING DATE: 2000-02-02
;  PRIOR APPLICATION NUMBER: 60/180921
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: 60/187202
;  PRIOR FILING DATE: 2000-03-03
;  PRIOR APPLICATION NUMBER: 60/198587
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: 60/199614
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 60/206330
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/206368
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/209832
;  PRIOR FILING DATE: 2000-06-05
;  PRIOR APPLICATION NUMBER: 60/218371
;  PRIOR FILING DATE: 2000-07-13
;  PRIOR APPLICATION NUMBER: 60/222695
;  PRIOR FILING DATE: 2000-08-02
;  PRIOR APPLICATION NUMBER: 60/229896
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/230621
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/232887
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: 60/235147
;  PRIOR FILING DATE: 2000-09-22
;  PRIOR APPLICATION NUMBER: 60/261878
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: 60/261910
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/261939
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/262150
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/264395
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/266421
;  PRIOR FILING DATE: 2001-02-02
;  PRIOR APPLICATION NUMBER: 60/267623
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/274399
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 60/280982
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: 60/282129
```

```
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: 60/282199
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: 60/290589
;  PRIOR FILING DATE: 2001-05-09
;  PRIOR APPLICATION NUMBER: 09/180997
;  PRIOR FILING DATE: 1998-11-19
;  PRIOR APPLICATION NUMBER: 09/267213
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 09/380137
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/380138
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/403297
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 09/423741
;  PRIOR FILING DATE: 1999-11-10
;  PRIOR APPLICATION NUMBER: 09/709238
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 09/802706
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 09/872035
;  PRIOR FILING DATE: 2001-06-01
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/924419
;  PRIOR FILING DATE: 2001-08-06
;  PRIOR APPLICATION NUMBER: 09/927796
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 09/929404
;  PRIOR FILING DATE: 2001-08-13
;  PRIOR APPLICATION NUMBER: 09/931836
;  PRIOR FILING DATE: 2001-08-16
;  PRIOR APPLICATION NUMBER: 09/941992
;  PRIOR FILING DATE: 2001-08-28
;  PRIOR APPLICATION NUMBER: 09/946374
;  PRIOR FILING DATE: 2001-09-04
;  PRIOR APPLICATION NUMBER: 10/001054
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 10/081056
;  PRIOR FILING DATE: 2002-02-20
;  PRIOR APPLICATION NUMBER: 10/119480
;  PRIOR FILING DATE: 2002-04-09

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy      507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
            ||  |||     |  | | || |||     |||   ||||||||  | ||  |   | | ||
Db      798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy      567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
              |    ||||||||       || || || || || || ||     | |||| ||| |
Db      858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy      627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
            ||  |||||||| ||||  | |||||||||||  |  | |  ||||||  |||| | |||||
Db      918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy      687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
            |||||| || ||| || | |||| ||||||  ||||||||||||| ||||| |     ||
Db      978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy      747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
              ||||    |||||   |  ||| ||  ||||||| || ||  |||| ||||| |
Db     1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy      807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
             || || ||||| |   || || ||||||||||||| |||  ||||| ||| | | ||||||
Db     1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy      867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
            ||||||||  |   | || |  |  ||  | || || |        |||| ||||||| ||
Db     1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy      927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
             ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db     1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy      987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
            | || ||  | || |||| |||||||||||||||||| |    |  ||   ||| | ||| |
Db     1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy     1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
               |||||| || ||||||| ||||||  | |||||||| || |||   |   | || ||
Db     1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy     1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
             || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db     1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy     1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
            | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db     1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy     1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
              | || | |||||| || ||||| ||  ||||| |  ||||||   |  ||    || ||
Db     1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy     1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
            ||||||||||   ||| | |||||||||||||| |||||||||||  |  | || || |
Db     1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy     1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
            | ||| |    ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db     1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy     1407 GGACCCAGCCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
            ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db     1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy     1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
              |  ||   | ||    ||||||| || ||||||  |  |   |||||    |    ||  |
Db     1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy     1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
            |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db     1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy     1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
              |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db     1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy     1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
               ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db     1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy     1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
             |      | | ||| |        | || || |  || |  ||| ||||        ||
Db     1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy     1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
            ||  | || || ||| |||| | |       | |    ||  |  |  |    |||||||||
Db     2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy     1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
             | ||||||||| |||| |  |||||  | ||  | |||||||||||||||| ||    | |
Db     2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy     1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db     2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226
```

```
Qy        1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
                |||  || ||||  ||||||   |||||||||||       |||||||||||||  |||
Db        2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy        2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
               | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db        2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy        2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
               |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db        2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy        2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
                |   |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db        2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy        2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
               || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db        2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy        2239 GCCGGGGCCCCGCCGCAGG 2257
                |||   |   |||   ||
Db        2527 CCCGCCCCGGAGCCCGCGG 2545


RESULT 10
US-10-238-183-99
; Sequence 99, Application US/10238183
; Publication No. US20030073189A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C11
;  CURRENT APPLICATION NUMBER: US/10/238,183
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/091358
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091978
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/099803
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/106932
;  PRIOR FILING DATE: 1998-11-03
;  PRIOR APPLICATION NUMBER: 60/115554
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/119342
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: 60/123957
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123972
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: 60/127372
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/131271
;  PRIOR FILING DATE: 1999-04-27
;  PRIOR APPLICATION NUMBER: 60/133459
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/135725
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/135729
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/135750
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/138385
;  PRIOR FILING DATE: 1999-06-09
;  PRIOR APPLICATION NUMBER: 60/140653
;  PRIOR FILING DATE: 1999-06-22
;  PRIOR APPLICATION NUMBER: 60/141037
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: 60/144732
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/144758
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/144790
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/145228
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/145698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: 60/146222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: 60/146843
;  PRIOR FILING DATE: 1999-08-03
;  PRIOR APPLICATION NUMBER: 60/148188
;  PRIOR FILING DATE: 1999-08-10
;  PRIOR APPLICATION NUMBER: 60/148513
;  PRIOR FILING DATE: 1999-08-12
;  PRIOR APPLICATION NUMBER: 60/149327
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/149395
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/150114
;  PRIOR FILING DATE: 1999-08-20
;  PRIOR APPLICATION NUMBER: 60/151700
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: 60/151734
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: 60/162506
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: 60/170262
;  PRIOR FILING DATE: 1999-12-09
;  PRIOR APPLICATION NUMBER: 60/177118
;  PRIOR FILING DATE: 2000-01-20
;  PRIOR APPLICATION NUMBER: 60/179851
;  PRIOR FILING DATE: 2000-02-02
;  PRIOR APPLICATION NUMBER: 60/180921
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: 60/187202
;  PRIOR FILING DATE: 2000-03-03
;  PRIOR APPLICATION NUMBER: 60/198587
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: 60/199614
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 60/206330
```

```
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/206368
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/209832
;  PRIOR FILING DATE: 2000-06-05
;  PRIOR APPLICATION NUMBER: 60/218371
;  PRIOR FILING DATE: 2000-07-13
;  PRIOR APPLICATION NUMBER: 60/222695
;  PRIOR FILING DATE: 2000-08-02
;  PRIOR APPLICATION NUMBER: 60/229896
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/230621
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/232887
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: 60/235147
;  PRIOR FILING DATE: 2000-09-22
;  PRIOR APPLICATION NUMBER: 60/261878
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: 60/261910
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/261939
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/262150
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/264395
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/266421
;  PRIOR FILING DATE: 2001-02-02
;  PRIOR APPLICATION NUMBER: 60/267623
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/274399
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 60/280982
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: 60/282129
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: 60/282199
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: 60/290589
;  PRIOR FILING DATE: 2001-05-09
;  PRIOR APPLICATION NUMBER: 09/180997
;  PRIOR FILING DATE: 1998-11-19
;  PRIOR APPLICATION NUMBER: 09/267213
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 09/380137
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/380138
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/403297
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 09/423741
;  PRIOR FILING DATE: 1999-11-10
;  PRIOR APPLICATION NUMBER: 09/709238
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 09/802706
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 09/872035
;  PRIOR FILING DATE: 2001-06-01
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/924419
;  PRIOR FILING DATE: 2001-08-06
;  PRIOR APPLICATION NUMBER: 09/927796
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 09/929404
;  PRIOR FILING DATE: 2001-08-13
;  PRIOR APPLICATION NUMBER: 09/931836
;  PRIOR FILING DATE: 2001-08-16
;  PRIOR APPLICATION NUMBER: 09/941992
;  PRIOR FILING DATE: 2001-08-28
;  PRIOR APPLICATION NUMBER: 09/946374
;  PRIOR FILING DATE: 2001-09-04
```

```
;  PRIOR APPLICATION NUMBER: 10/001054
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 10/081056
;  PRIOR FILING DATE: 2002-02-20
;  PRIOR APPLICATION NUMBER: 10/119480
;  PRIOR FILING DATE: 2002-04-09

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
              ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db        798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                |    ||||||||       || || || || || || ||     | ||||| ||| |
Db        858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
              ||  |||||||| ||||  | |||||||||||  |  | |  ||||||  ||| | |||||
Db        918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
              |||||| || ||| || | ||| ||||||  ||||||||||||| ||||| |     ||  
Db        978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |   
Db       1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
               || || ||||| |   || || |||||||||||||| |||  ||||| ||| | | ||||||
Db       1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
              ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db       1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
               ||||||||||||||||||||| |    |||| |||||  || || | ||| ||| |||
Db       1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
              | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db       1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                 |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db       1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
               || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db       1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
              | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db       1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                | || | |||||| || |||| ||  |||| |  ||||||   |  ||    || ||
Db       1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
              ||||||||||   ||| | ||||||||||||| |||||||||||  |  | || || | 
Db       1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
              | |||       ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db       1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694
```

```
Qy        1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
               ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db        1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy        1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                 |  ||   | ||    |||||| || ||||  |  |  |||||    |    ||  |
Db        1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy        1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
               |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db        1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy        1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                 |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db        1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy        1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                  ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db        1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy        1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
                |      | | ||| |        | || || |  || |  ||| ||||        ||
Db        1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy        1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
               ||  | || || ||| ||||| | | |       | |    ||  |  |    |||||||||
Db        2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy        1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
                | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db        2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy        1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                   | |||  || | |   ||||| |||||      ||||||  |||| ||||| |||
Db        2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy        1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
                |||  || ||||  ||||||   ||||||||||||       ||||||||||||||  |||
Db        2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy        2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
               | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db        2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy        2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
               |||||| | |   | || || ||||||| ||| || | |||||||||||| || | ||
Db        2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy        2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
                |  |||||||||||||   |  |  | || |||| |       ||||||  | ||| |
Db        2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy        2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
               || ||||  || | ||| ||||||||||||||||| |   ||  | || |    ||| ||
Db        2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy        2239 GCCGGGGCCCCGCCGCAGG 2257
                |||   |   |||   ||
Db        2527 CCCGCCCCGGAGCCCGCGG 2545

RESULT 11
US-10-238-283-99
; Sequence 99, Application US/10238283
; Publication No. US20030073190A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C15
;  CURRENT APPLICATION NUMBER: US/10/238,283
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-238-283-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy         507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
               ||  |||     |  | | || |||     |||   ||||||||  | ||   | | ||
Db         798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy         567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                 |    ||||||||       || || || || || || ||     | ||||| ||| |
Db         858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy         627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
               ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db         918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy         687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
               |||||| || ||| || | ||| ||||||  ||||||||||||| ||||| |     ||
Db         978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy         747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                 ||||    |||||   |  ||| ||  |||||||| || ||  |||| ||||| |
Db        1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy         807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
                || || ||||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db        1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy         867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
               ||||||||  |   | || |  |  ||  | || ||  |      |||| ||||||| ||
Db        1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy         927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
                ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db        1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277
```

```
Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
              | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db       1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                 |||||| || ||||||||| |||||  | |||||||| || |||   |   | || ||
Db       1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
               || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db       1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
              | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db       1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db       1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
              ||||||||||   ||| | |||||||||||||| |||||||||||  |  | || || |
Db       1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
              | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db       1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
              ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db       1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                |  ||   | ||    |||||| || |||||  |  |  |||||    |    ||  |
Db       1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
              |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db       1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db       1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                 ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db       1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
               |      | | ||| |        | || || |  || |  ||| ||||        ||
Db       1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
              ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db       2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
               | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db       2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                  | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db       2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
               |||  || ||||  ||||||   |||||||||||       |||||||||||||  |||
Db       2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
              | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db       2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
              |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db       2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
               |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db       2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
              || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db       2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257
               |||   |   |||   ||
Db       2527 CCCGCCCCGGAGCCCGCGG 2545


RESULT 12
US-10-238-370-99
; Sequence 99, Application US/10238370
; Publication No. US20030073191A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C10
;  CURRENT APPLICATION NUMBER: US/10/238,370
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-238-370-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
              ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
```

```
Db         798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy         567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                 |    ||||||||       || || || || || || ||     | ||||| ||| |
Db         858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy         627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
               ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db         918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy         687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
               |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db         978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy         747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                 ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db        1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy         807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
                || || |||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db        1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy         867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
               ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db        1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy         927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
                ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db        1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy         987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
               | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db        1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy        1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                  |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db        1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy        1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
                || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db        1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy        1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
               | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db        1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy        1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                 | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db        1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy        1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
               ||||||||||   ||| | |||||||||||||| |||||||||||  |  | || || |
Db        1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy        1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
               | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db        1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy        1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
               ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db        1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy        1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                 |  ||   | ||    |||||| || |||||  |  |  |||||    |    ||  |
Db        1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy        1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
               |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db        1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy        1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                 |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db        1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy        1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                  ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db        1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy        1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
                |      | | ||| |        | || || |  || |  ||| ||||        ||
Db        1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy        1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
               ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db        2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy        1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
                | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db        2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy        1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                   | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db        2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy        1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
                |||  || ||||  ||||||   |||||||||||       |||||||||||||  |||
Db        2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy        2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
               | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db        2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy        2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
               |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db        2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy        2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
                |  |||||||||||||   |  |  | || ||||  |      ||||||  | ||| |
Db        2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCGGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy        2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
               || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db        2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy        2239 GCCGGGGCCCCGCCGCAGG 2257
                |||   |   |||   ||
Db        2527 CCCGCCCCGGAGCCCGCGG 2545
```

```
RESULT 13
US-10-245-055-99
; Sequence 99, Application US/10245055
; Publication No. US20030073192A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C88
;  CURRENT APPLICATION NUMBER: US/10/245,055
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
```

```
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-055-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566

Db        798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626

Db        858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686

Db        918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746

Db        978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806

Db       1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866

Db       1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926

Db       1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986

Db       1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046

Db       1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106

Db       1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166

Db       1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226

Db       1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286

Db       1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346

Db       1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406

Db       1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466

Db       1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526

Db       1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586

Db       1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644

Db       1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704

Db       1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764

Db       1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824

Db       2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884

Db       2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944

Db       2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004

Db       2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064

Db       2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124

Db       2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178

Db       2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238

Db       2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGCCCCACGCCCACGCCCGCCCCG 2526

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257

Db       2527 CCCGCCCCGGAGCCCGCGG 2545
```

```
RESULT 14
US-10-245-147-99
; Sequence 99, Application US/10245147
; Publication No. US20030073193A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C72
;  CURRENT APPLICATION NUMBER: US/10/245,147
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-147-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
              ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db        798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                |    ||||||||       || || || || || || ||     | ||||| ||| |
Db        858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
              ||  |||||||| ||||  | ||||||||||  |  | |  ||||||  ||| | |||||
Db        918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
              |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db        978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db       1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
               || || |||| |   || || |||||||||||| |||  ||||| ||| | | ||||||
Db       1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
              ||||||||  |   | || |  |  ||  | || || |       |||| ||||||| ||
Db       1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
               ||||||||||||||||||||| |    ||||| |||||  || || | ||| ||| |||
Db       1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
              | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db       1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                 |||||| || |||||||| |||||  | |||||||| || |||   |   | || ||
Db       1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
               || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db       1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
              | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db       1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                | || | |||||| || ||||| ||  |||| |  ||||||   |  ||    || ||
Db       1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
              ||||||||||   ||| | |||||||||||||| |||||||||||  |  | || || |
Db       1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
              | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db       1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
              ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db       1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy       1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                |  ||   | ||    |||||| || |||||  |  |  |||||    |    ||  |
Db       1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy       1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
              |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db       1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy       1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db       1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy       1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                 ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db       1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy       1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
               |      | | ||| |        | || || |  || |  ||| ||||        ||
Db       1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy       1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
              ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db       2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy       1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
               | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
```

```
Db       2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy       1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                  | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db       2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy       1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
               |||  || ||||  ||||||   |||||||||||       ||||||||||||||  |||
Db       2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy       2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
              | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db       2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy       2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
              |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db       2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy       2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
               |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db       2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy       2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
              || ||||  || | ||| ||||||||||||||| |   ||  | || |    ||| ||
Db       2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy       2239 GCCGGGGCCCCGCCGCAGG 2257
               |||   |   |||   ||
Db       2527 CCCGCCCCGGAGCCCGCGG 2545


RESULT 15
US-10-245-730-99
; Sequence 99, Application US/10245730
; Publication No. US20030073194A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C85
;  CURRENT APPLICATION NUMBER: US/10/245,730
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
;   LENGTH: 2855
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-245-730-99

  Query Match             19.8%;  Score 621.8;  DB 9;  Length 2855;
  Best Local Similarity   61.3%;  Pred. No. 1.3e-162;
  Matches 1078;  Conservative    0;  Mismatches  662;  Indels   19;  Gaps    4;

Qy        507 GTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566
              ||  |||     |  | | || |||     |||   ||||||||  | |||   | | ||
Db        798 GTCAGCCACACCCAGCCCATGTCCCCGCCGCTGCCGCTGCCAGACACAGTCGCTGCCCCT 857

Qy        567 GGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGAC 626
                |    ||||||||       || || || || || || ||     | ||||| ||| |
Db        858 AAGCGTGCTGTGCCCAGGGGCAGGCCTCCTGTTCGTGCCACCCTCGCTGGACCGCCGGGC 917

Qy        627 AGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAA 686
              ||  |||||||| ||||  | |||||||||||  |  | |  ||||||  ||| | |||||
Db        918 AGCCGAGCTGCGGCTGGCAGACAACTTCATCGCCTCCGTGCGCCGCCGCGACCTGGCCAA 977

Qy        687 CATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
              |||||| || ||| || | |||| ||||||  ||||||||||| ||||| |     ||
Db        978 CATGACAGGCCTGCTGCATCTGAGCCTGTCGCGGAACACCATCCGCCACGTGGCTGCCGG 1037

Qy        747 TTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAG 806
                ||||    |||||   |  ||| ||  ||||||| || ||  |||| |||||| |
Db       1038 CGCCTTCGCCGACCTGCGGGCCCTGCGTGCCCTGCACCTGGATGGCAACCGGCTGACCTC 1097

Qy        807 CCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAA 866
               || || ||||| |   || || ||||||||||||| |||  ||||| ||| | | ||||||
Db       1098 ACTGGGCGAGGGCCAGCTGCGCGGCCTGGTCAACTTGCGCCACCTCATCCTCAGCAACAA 1157

Qy        867 CCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCT 926
              ||||||||  |   | || |  |  ||  | || || |  |       |||| ||||||| ||
Db       1158 CCAGCTGGCAGCGCTGGCGGCCGGCGCCCTGGATGATTGTGCCGAGACACTGGAGGACCT 1217

Qy        927 GGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAA 986
               ||||||||||||||||||||| |    |||| |||||  || || | ||| ||| |||
Db       1218 CGACCTCTCCTACAACAACCTCGAGCAGCTGCCCTGGGAGGCCCTGGGCCGCCTGGGCAA 1277

Qy        987 CCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGC 1046
              | || ||  | || |||| |||||||||||||||| |    |  ||   ||| | ||| |
Db       1278 CGTCAACACGTTGGGCCTCGACCACAACCTGCTGGCTTCTGTGCCC---GGCGCTTTTTC 1334

Qy       1047 AGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCC 1106
                 |||||| || ||||||||| |||||  | |||||||| || |||   |   | || ||
Db       1335 CCGCCTGCACAAGCTGGCCCGGCTGGACATGACCTCCAACCGCCTGACCACAATCCCACC 1394

Qy       1107 TGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACC 1166
               || ||  ||||  |||||   |        ||   | |        |    ||| |  |
Db       1395 CGACCCACTCTTCTCCCGCCTGCCCCTGCTCGCCAGGCCCCGGGGCTCGCCCGCCTCTGC 1454

Qy       1167 CTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCG 1226
              | ||    |    ||||| || ||||| || |||||||| || |||||  | ||||||||
Db       1455 CCTGGTGCTGGCCTTTGGCGGGAACCCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGCG 1514

Qy       1227 GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCG 1286
                | || | ||||||| || ||||| ||  ||||| |  ||||||   |  ||    || ||
Db       1515 TCGCCTGGCGCGGGAGGACGACCTCGAGGCCTGCGCGTCCCCACCTGCTCTGGGCGGCCG 1574

Qy       1287 CTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCA 1346
              ||||||||||   ||| | |||||||||||||| ||||||||||||  |  | || || |
Db       1575 CTACTTCTGGGCGGTGGGCGAGGAGGAGTTTGTCTGCGAGCCGCCCGTGGTGACTCACCG 1634

Qy       1347 CACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGG 1406
              | |||      ||   || |  |  || | | ||||  | ||    |||   ||  | ||
Db       1635 CTCACCACCTCTGGCTGTGCCCGCAGGTCGGCCGGCTGCCCTGCGCTGCCGGGCAGTGGG 1694

Qy       1407 GGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTC 1466
```

```
                 ||||||    |||| | | |  |||||  | ||| | | ||| ||| |||| |||||
Db          1695 GGACCCAGAGCCCCGTGTGCGTTGGGTGTCACCCCAGGGCCGGCTGCTAGGCAACTCAAG 1754

Qy          1467 AAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAG 1526
                   |  ||   | ||    |||||| || |||||  |  |  |||||    |    ||  |
Db          1755 CCGTGCCCGCGCCTTCCCCAATGGGACGCTGGAGCTGCTGGTCACCGAGCCGGGTGATGG 1814

Qy          1527 TGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGT 1586
                 |||   ||||||||||||||| |||||||| || || |||||||| ||    |||||| |
Db          1815 TGGCATCTTCACCTGCATTGCGGCCAATGCAGCTGGCGAGGCCACAGCTGCTGTGGAGCT 1874

Qy          1587 CTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGCACCCCCCAAG 1644
                   |  | |  |  |  ||||  |          || ||||  ||      |||||||  |
Db          1875 GACTGTGGGTCCCCCACCACCTCCTCAGCTAGCCAACAGCACCAGCTGTGACCCCCCGCG 1934

Qy          1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGG 1704
                    ||    ||      | ||  | ||| |     | |||   |     |     ||||
Db          1935 GGACGGGGATCCTGATGCTCTCACCCCACCCTCCGCTGCCTCTGCTTCTGCCAAGGTGGC 1994

Qy          1705 GGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACC 1764
                  |      | | ||| |        | || || |  || |  ||| ||||        ||
Db          1995 CGACACTGGGCCCCCTA--------CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046

Qy          1765 ACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAG 1824
                 ||  | || || ||| ||||| | |       | |    ||  |  |    |||||||||
Db          2047 ACAGCTGCTCTTGTCCAGTGGCCGGATCAGCGGCCTATCCCGGGCATCCGCATGTACCAG 2106

Qy          1825 CTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAAC 1884
                  | ||||||||| |||| |  |||||  | ||  | |||||||||||||| ||    | |
Db          2107 ATCCAGTACAACAGCTCGGCTGATGACATCCTCGTCTACAGGATGATCCCGGCGGAGAGC 2166

Qy          1885 AAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCC 1944
                     | ||| || | |   |||||| ||||||      ||||||  |||| ||||| |||
Db          2167 CGCTCGTTCCTGCTGACGGACCTGGCGTCAGGCCGGACCTACGATCTGTGCGTGCTCGCC 2226

Qy          1945 ATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGCTGCGCCCAGTTC 2004
                  |||  || ||||  ||||||   |||||||||||       |||||||||||||  |||
Db          2227 GTGTATGAGGACAGCGCCACGGGGCTCACGGCCACGCGGCCTGTGGGCTGCGCCCGCTTC 2286

Qy          2005 TTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGATTCTGGGCGGCACC 2064
                 | |||| |  |||     | ||  |||  | |   ||||   |   | |||||||||||
Db          2287 TCCACCGAACCTGCGCTGCGGCCATGCGGGGCGCCGCACGCTCCCTTCCTGGGCGGCACG 2346

Qy          2065 ATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCTTCATCGTCATCCTC 2124
                 |||||| | |   | || ||| ||||||| ||| || | |||||||||||| || | ||
Db          2347 ATGATCATCGCGCTGGGCGGCGTCATCGTAGCCTCGGTACTGGTCTTCATCTTCGTGCTG 2406

Qy          2125 ATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAG------CAAGATGGCAGCGGCC 2178
                  |  |||||||||||||   |  |  | || ||||| |      ||||||  | ||| |
Db          2407 CTAATGCGCTACAAGGTGCACGGCGGCCAGCCCCCCGGCAAGGCCAAGATTCCCGCGCCT 2466

Qy          2179 GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAAGCAGCGCACCA 2238
                 || ||||  || | ||| |||||||||||||||| |   ||  | || |    ||| ||
Db          2467 GTTAGCAGCGTTTGCTCCCAGACCAACGGCGCCCTGGGCCCCACGCCCACGCCCGCCCCG 2526

Qy          2239 GCCGGGGCCCCGCCGCAGG 2257
                  |||   |   |||   ||
Db          2527 CCCGCCCCGGAGCCCGCGG 2545

Search completed: June  8, 2003, 21:27:39
Job time : 453 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         June  8, 2003, 15:32:52 ; Search time 3910 Seconds
                                          (without alignments)
                                          13022.662 Million cell updates/sec

Title:          US-09-831-846-1
Perfect score:  3144
Sequence:       1 gcctggctccctctcgctga..........gcggctttggattgcttatg 3144

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  gb_gss:*
                18:  em_gss_hum:*
                19:  em_gss_inv:*
                20:  em_gss_pln:*
                21:  em_gss_vrt:*
                22:  em_gss_fun:*
                23:  em_gss_mam:*
                24:  em_gss_mus:*
                25:  em_gss_other:*
                26:  em_gss_pro:*
                27:  em_gss_rod:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                      %
Result              Query
   No.     Score   Match Length DB  ID                         Description
------------------------------------------------------------------------------
     1    2170.6    69.0   3204  11  AK017594                  AK017594 Mus muscu
     2     609.4    19.4    611  14  BM787163                  BM787163 K-EST0066
     3     606.8    19.3    621  12  BF312768                  BF312768 601896906
     4       582    18.5    582  14  BQ126354                  BQ126354 ii16e11.y
c    5     540.2    17.2    578  14  BQ126066                  BQ126066 ii16e11.x
     6     525.6    16.7    652  14  BM948451                  BM948451 UI-M-EG0p
     7     514.2    16.4    655  10  BB642310                  BB642310 BB642310
     8     492.6    15.7   1853  11  BC007718                  BC007718 Homo sapi
     9     441.8    14.1    457  10  AW163387                  AW163387 au94f03.y
    10     422.8    13.4    435  10  AW161566                  AW161566 au71e01.y
    11     377.2    12.0    685  10  BB278209                  BB278209 BB278209
    12     346.6    11.0    357  13  BM193100                  BM193100 TCBAP1D95
c   13       327    10.4    635  12  BG380772                  BG380772 UI-R-CT0-
c   14       313    10.0    397  10  AW161909                  AW161909 au71e01.x
    15     310.6     9.9    708  13  BM046297                  BM046297 603626095
    16     300.4     9.6    714  14  BM950379                  BM950379 UI-M-EH0p
    17     293.8     9.3    674  14  BM722122                  BM722122 UI-E-EO0-
    18     285.4     9.1    338  10  AW415047                  AW415047 49056 MAR
    19       279     8.9    563  10  AW274912                  AW274912 xm62f02.x
    20     276.6     8.8    712  14  BM944080                  BM944080 UI-M-EH0p
    21     273.4     8.7    822  14  BQ571534                  BQ571534 UI-M-FC0-
    22     271.4     8.6    703  14  BM784343                  BM784343 K-EST0062
    23     267.4     8.5    472  12  BF563759                  BF563759 UI-R-BT1-
    24     266.2     8.5    512   9  AI383513                 AI383513 td03d01.x
    25     266.2     8.5    513   9  AI418838                 AI418838 tf37e04.x
    26     264.8     8.4    526  10  AW003909                  AW003909 wq84a05.x
    27     264.4     8.4    720  14  BM948138                  BM948138 UI-M-EG0p
    28     263.2     8.4    635  14  BM742719                  BM742719 K-EST0015
    29       263     8.4    513   9  AI760811                 AI760811 wi67g10.x
c   30     260.6     8.3    634   9  AI830569                 AI830569 wj51c10.x
    31     259.8     8.3    618  14  BM744495                  BM744495 K-EST0018
    32     259.8     8.3    620  14  BM742162                  BM742162 K-EST0014
    33     259.8     8.3    620  14  BM742315                  BM742315 K-EST0015
    34       259     8.2   1091  13  BM552976                  BM552976 AGENCOURT
    35     258.2     8.2    609  14  BM742747                  BM742747 K-EST0015
    36     257.6     8.2    609  14  BM742740                  BM742740 K-EST0015
    37     257.4     8.2    751  12  BE879258                  BE879258 601484911
    38     252.4     8.0    602  14  BM741814                  BM741814 K-EST0014
    39       252     8.0    618  14  BM744466                  BM744466 K-EST0018
c   40     249.6     7.9    541  12  BF411974                  BF411974 UI-R-BT1-
    41     249.6     7.9    589  14  BM741945                  BM741945 K-EST0014
    42     245.4     7.8    600  14  BM949682                  BM949682 UI-M-EG0p
    43     243.8     7.8    581  14  BM783370                  BM783370 K-EST0061
    44     243.6     7.7    526   9  AA527134                 AA527134 ni07a02.s
    45       243     7.7    637  14  BM784328                  BM784328 K-EST0062


                                 ALIGNMENTS


RESULT 1
AK017594
LOCUS       AK017594                3204 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 8 days embryo whole body cDNA, RIKEN full-length
            enriched library, clone:5730420O05:homolog to KIAA1246 PROTEIN
            (FRAGMENT), full insert sequence.
ACCESSION   AK017594
VERSION     AK017594.1  GI:12856913
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 8 days embryo cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:5730420O05.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
```

```
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 3204)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of subtraction to Rot = 100.0. Second strand cDNA
            was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..3204
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:5730420O05"
                     /db_xref="MGD:MGI:1902927"
                     /db_xref="taxon:10090"
                     /clone="5730420O05"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="8 days embryo"
     CDS             466..2832
                     /note="data source:SPTR, source key:Q9ULH4, evidence:ISS
                     homolog to KIAA1246 PROTEIN (FRAGMENT)
                     putative"
                     /codon_start=1
                     /protein_id="BAB30828.1"
                     /db_xref="GI:12856914"
                     /db_xref="MGD:MGI:1917780"
                     /translation="METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKRLL
                     FVPPDIDRRTVELRLGGNFIIHIGRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESL
                     RSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADDAFEDFLLTLEDLDLSYNN
                     LHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIF
                     ARSQASLLTATPFAPPLSFSFGGNPLHCNCELLWLRRLERDDDLETCGSPGSLKGRYF
                     WHIREEEFVCEPPLITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSS
                     RTAVYDNGTLDILITTSQDSGPFTCIAANAAGEATATVEVSIVQLPHLSNSTSRMAPP
                     KSRLSDITGSSKTSRGGGSGAGEPPKSTPERAVLVSDVTTTSALVKWSVSKSAPRVK
                     MYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIV
                     GCAQFFTKADYPQCQSMHSQIKGGTMILVIGGIIVATLLVFIVILMVRYKVCNHDTPG
                     KMAAATVSNVYSQTNGSQPPPLGGIPVGQLPQAPPKVVVRNELMDFSTSLARACDSSS
                     SSSLGSGEAAGLGRGPWRLPPPAPRPKPSLDRLMGAFASLDLKSQRKEELLDSRTPAG
                     RGAGTSSRGHHSDREPLLGPPATRARSLLPLPLEGKAKRSHSFDMGDFAAAAAAVPGG
                     YSPPRRVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSEWVMESTV"
BASE COUNT      660 a    964 c    864 g    716 t
ORIGIN

  Query Match             69.0%;  Score 2170.6;  DB 11;  Length 3204;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches 2651;  Conservative    0;  Mismatches  439;  Indels   67;  Gaps   15;

Qy         47 ACCCGGCAGGCTCGTCTGAACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTG 106
              || || ||| || ||||||||||||||||| |||||||||||||| ||   |||||||||
Db         11 ACACGACAGACTGGTCTGAACTTGAAGACATCCCACATTCCAAGAGGCTGCAGGTTCCTG 70

Qy        107 GGAATGCC--TGGGGTTCTTCGATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGAT 164
              || |||||   |||||||||| |   |||||||||||||  |||| ||||||   ||  |
Db         71 GGGATGCCAGGGGGGTTCTTCAACAAGGAAAATCCTACCCTCATCTTCCTAGAAGGGATT 130

Qy        165 TATTATT----------------ATTATTTTTCTTTAATCTGGAAGAGAAGAGAACAAGT 208
                 | ||                 |||||||| |  ||||||||||||||||| ||||||
Db        131 ATCTTTTTTTATCCTTTTTCCCCCTTATTTTTTTAAAATCTGGAAGAGAAGAGCACAAGT 190

Qy        209 TGTGCT-----TTTCCCCCCTTCTTCTTGCTAAACGCCATGGATATAACTGAATAAGCGG 263
              ||| ||     |  | |||||||||||||||||| ||||||||||||||||||||||| |
Db        191 TGTACTCCCCCTCCCGCCCCTTCTTCTTGCTAAATGCCATGGATATAACTGAATAAGCTG 250

Qy        264 CTCAGGGCTTTCCCCGCGTGGA-CGTCCGAGGCCACCATCTGCCTGCATTCGCCGGA--- 319
              ||||||||| ||||||||||||  ||||||||||||||||||||||| ||||| |||
Db        251 CTCAGGGCTCTCCCCGCGTGGATGGTCCGAGGCCACCATCTGCCTGCCTTCGCTGGATCG 310

Qy        320 --GCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCC 377
                || || |||| | |||||  ||| ||||||| |||  |||||||||||||||||||||
Db        311 CTGCAGCTGGAGCGCTTAGCGTGAGCCTGTCTCCGGCCAGGAAGGATGCGTGGCCGAGCC 370

Qy        378 GGGGAGCCCGGGCGCCCCGCGGAGCC-GGCCTCGGTGCCACCCAGCCGGGGGTAGATGCT 436
                ||||||||||||||| |||||||| ||||| || ||| ||| |||  ||||||||||
Db        371 AAGGAGCCCGGGCGCCCTGCGGAGCCGGGCCTTGGAGCCTCCCCGCCACGGGTAGATGCC 430

Qy        437 GCCTCGCCCAGG------CGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGGCCTGC 490
              |||| |||||||       |||||||||||||||||||||||| ||||||||||| ||||
Db        431 GCCTTGCCCAGGTGCAGAGGCTGAGTGACCAGACCATGGAGACTCTGCTTGGTGGGCTGC 490
```

```
Qy    491 TAGCGTTTGGCATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGA 550
          | || ||||||||||||||||| ||||||||| |||||||||||||||||| ||||||||||
Db    491 TGGCTTTTGGCATGGCGTTTGCTGTGGTCGATGCCTGCCCCAAGTACTGCGTCTGCCAGA 550

Qy    551 ATCTGTCTGAGTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTG 610
          |||||||||||||||||||||||||||| ||||||||| |||| |||||||| |||||||
Db    551 ATCTGTCTGAGTCACTGGGGACCCTGTGTCCCTCCAAGAGGCTCCTCTTTGTGCCCCCTG 610

Qy    611 ATATTGACCGGCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCC 670
          | ||||||||| |||| |||||||||||| ||||| ||||||||||||||||||||  |||
Db    611 ACATTGACCGGAGGACGGTGGAGCTGCGTCTGGGTGGCAACTTCATCATCCACATTGGCC 670

Qy    671 GCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCA 730
          |||| ||||||||||||||||| |||||||||||| |||||| |||||||||||||||||||
Db    671 GCCAAGACTTTGCCAACATGACAGGGCTGGTGGATCTGACCTTGTCCAGGAACACCATCA 730

Qy    731 GCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACA 790
          |||| ||||||||||| ||||| ||||||||| |||||||||||||||||||| |||||||
Db    731 GCCATATCCAGCCCTTCTCCTTCCTGGACCTTGAGAGCCTCCGCTCCCTGCACCTTGACA 790

Qy    791 GCAATCGGCT GCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACC 850
```

```
Qy    791 GCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACC 850
          |||| ||||| || |||||||||||||||||| ||||||||| |||||||||||||| ||||
Db    791 GCAACCGGCTACCTAGCCTTGGGGAGGACACACTCCGGGGTCTGGTCAACCTGCAACACC 850

Qy    851 TTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGC 910
          |||| |||||||| |||||||||| |||||||| ||||| || ||||| |||||| |||
Db    851 TTATTGTGAACAATAACCAGCTGGGTGGCATCGCTGATGATGCCTTTGAAGACTTCTTGC 910

Qy    911 TGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCG 970
          |||| || || || |||||||| || |||||||||||||||| || || || ||||| ||||
Db    911 TGACTTTAGAAGACCTGGACCTATCTTACAACAACCTCCACGGACTTCCCTGGGATTCCG 970

Qy    971 TGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCG 1030
          | || |||||||||||||||| ||||||||| ||||||||||| |||||||| ||||| |
Db    971 TACGGCGCATGGTCAACCTCCATCAGCTGAGTCTGGACCACAATCTGCTGGACCACATTG 1030

Qy   1031 CCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGC 1090
          | |||||||| |||||||||||||||||||||||||||||||| ||||| ||||||||||
Db   1031 CTGAGGGCACTTTTGCAGACCTGCAGAAACTGGCCCGCCTGGACCTCACATCCAATCGGC 1090

Qy   1091 TGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCA 1150
          ||||||| || ||||| || ||||||||||||||||||||||||||     | || ||||
Db   1091 TGCAGAAACTTCCCCCGGACCCCATCTTTGCCCGCTCCCAGGCTTCCTTGCTAACTGCCA 1150

Qy   1151 CACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTG 1210
          | ||||||||||||||| |||| |||||||||||| || ||||| || |||||||||||||
Db   1151 CGCCCTTTGCCCCACCCCTGTCTTTTAGTTTTGGAGGGAACCCGCTGCACTGCAATTGTG 1210

Qy   1211 AGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAG 1270
          |||||||||||||||||||||| ||| |||| |||||||||||| |||||||| ||||| |
Db   1211 AGCTTCTCTGGCTGCGGAGGCTGGAGAGGGATGATGACCTGGAGACCTGTGGATCCCCCG 1270

Qy   1271 GGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGC 1330
          |  | ||||||||||||||||| |||||| | ||||||||||||||||||||||| |||||||
Db   1271 GAAGTCTCAAGGGTCGCTACTTTTGGCATATTCGTGAGGAGGAGTTTGTGTGTGAGCCGC 1330

Qy   1331 CTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCA 1390
          |||||||||||||||||||||||||| ||||||||||||||||||||||| ||||| || |
Db   1331 CTCTCATCACCCAGCACACACACAAGCTGCTGGTTCTGGAGGGCCAGGCAGCCACTCTTA 1390

Qy   1391 AGTGCAAAGCCATTGGGGACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCC 1450
          ||||||| ||||||||||||||||||||  | ||||||||||| |||||||||||||||
Db   1391 AGTGCAAGGCCATTGGGGACCCCAGCCCTTTGATCCACTGGGTTGCCCCCGATGACCGCT 1450

Qy   1451 TGGTAGGGAACTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCA 1510
          |||| || ||||| || |||||| ||||| |||||||||||||||||||||||||  | || |
Db   1451 TGGTGGGAAACTCTTCCAGGACTGCTGTATATGACAATGGCACCCTGGACATTCTAATTA 1510

Qy   1511 CCACATCTCAGGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCA 1570
          |||| ||||||||||||||  |||| |||||||| || |||||||| || |||||||| |
Db   1511 CCACCTCTCAGGACAGTGGACCCTTTACCTGCATCGCAGCCAATGCAGCAGGAGAGGCTA 1570

Qy   1571 CGGCCATGGTGGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA 1630
          ||||||  || ||||||||||| || |||||| |||||||| || |||||||| ||||| |
Db   1571 CGGCCACTGTAGAGGTCTCCATTGTGCAGCTCCCACACCTTAGTAACAGCACTAGCCGGA 1630

Qy   1631 CTGCACCCCCCAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAG 1690
            |||||||||||||| ||||| || ||||||||| || ||||||||||||||||||||||
Db   1631 TGGCACCCCCCAAGTCTCGCCTTTCGGACATCACAGGTTCCAGCAAGACCAGCCGGGGAG 1690

Qy   1691 GTGGAGGCAGTGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGT 1750
          | ||||| ||||||||  || ||||||||||||||| |||| ||  |||||||||||||||
Db   1691 GGGGAGGTAGTGGGGCTGGGGAGCCTCCCAAAAGCACCCCAGAGAGGGCTGTGCTTGTGT 1750

Qy   1751 CTGAAGTGACCACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGG 1810
          | || || |||||||| || |||||||||||||||||||||||||||||||||||||| |||
Db   1751 CCGATGTTACCACCACATCTGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCAGGG 1810

Qy   1811 TGAAGATGTACCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGA 1870
          |||||||||| ||||||||||||||||| ||||| |||||||| ||||| ||||||||||
Db   1811 TGAAGATGTATCAGCTGCAGTACAACTGTTCTGATGATGAGGTGCTGATCTACAGGATGA 1870

Qy   1871 TCCCAGCCTCCAACAAGGCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACT 1930
          |||||||||| ||||||||||||||||||||||||||||||||| || || ||||||||||
Db   1871 TCCCAGCCTCTAACAAGGCCTTCGTGGTCAACAACCTGGTGTCCGGAACGGGCTACGACT 1930

Qy   1931 TGTGTGTGCTGGCCATGTGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGG 1990
          |||| || ||||| |||||||| ||||||||||||||| |||| |||||||||||| ||||
Db   1931 TGTGCGTACTGGCTATGTGGGACGACACAGCCACGACGCTCACAGCCACCAACATTGTGG 1990

Qy   1991 GCTGCGCCCAGTTCTTCACCAAGGCTGACTACCCGCAGTGCCAGTCCATGCACAGCCAGA 2050
          |||| |||||||||||||||||||| |||||||| || |||||||| |||||||| ||||
Db   1991 GCTGTGCCCAGTTCTTCACCAAGGCCGACTACCCACAATGCCAGTCTATGCACAGTCAGA 2050

Qy   2051 TTCTGGGCGGCACCATGATCCTGGTCATCGGGGGCATCATCGTGGCCACGCTGCTGGTCT 2110
          |   ||| ||||||||||| ||||||||||| || ||||| ||||||||  ||||||||||
Db   2051 TAAAGGGGGGCACCATGATTCTGGTCATCGGAGGTATCATTGTGGCCACACTGCTGGTCT 2110

Qy   2111 TCATCGTCATCCTCATGGTGCGCTACAAGGTCTGCAACCACGAGGCCCCCAGCAAGATGG 2170
          |||| || |||||||||||||| |||||||||||||||||||||  ||||  | || ||||
Db   2111 TCATTGTAATCCTCATGGTGCGGTACAAGGTCTGCAACCACGACACCCCAGGAAAAATGG 2170

Qy   2171 CAGCGGCC---GTGAGCAATGTGTACTCGCAGACCAACGGCGCCCAGCCACCGCCTCCAA 2227
          |||| |||    || |||||||||||||||||||||| ||  | |||||  || ||||
Db   2171 CAGCCGCCACTGTCAGCAATGTGTACTCGCAGACCAATGGGTCTCAGCCTCCACCTCTGG 2230

Qy   2228 GCAGCGCACCAGCCGGGGCCCCGCCGCAGGGCCCGCCGAAGGTGGTGGTGCGCAACGAGC 2287
          |  |   ||| |  ||    |  || ||||  || || || || |||||||| || |||
Db   2231 GTGGGATACCCGTTGGCCAGCTTCCTCAGGCACCACCCAAAGTTGTGGTGCGAAATGAGT 2290

Qy   2288 TCCTGGACTTCACCGCCAGCCTGGCCCGCGCCAGTGACTCCTCTTCCTCCAGCTCCCTGG 2347
          |  |||||||||   ||||||||||||| | ||| |||||||||||||| ||||||||| ||||
Db   2291 TGATGGACTTCAGTACCAGCCTGGCCAGGGCCTGTGACTCCTCTTCTTCCAGCTCACTGG 2350

Qy   2348 GCAGTGGGGAGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACCCTCCGCCCCGC 2407
          |||| ||||| || |||||||||||| ||||  |||||||||| ||||||||| | || || |
Db   2351 GCAGCGGGGAAGCCGCGGGGCTGGGCCGGGGTCCCTGGAGGCTCCCACCCCCAGCTCCTC 2410

Qy   2408 GCCCCAAGCCCAGCCTTGACCGCCTGATGGGGGCCTTCGCCTCCCTGGACCTCAAGAGTC 2467
          | |||||||||||||||||||||| ||||||| ||||| |||||||||||||||||||||||
Db   2411 GTCCCAAGCCCAGCCTTGACCGCTTGATGGGAGCCTTTGCCTCCCTGGACCTCAAGAGTC 2470

Qy   2468 AGAGAAAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGGACGTCGG 2527
          | || || || || || || |||||||||||||||||| || |||||||| ||||| ||
Db   2471 AAAGGAAAGAAGAACTTCTAGACTCCAGGACTCCAGCAGGCAGAGGGGCAGGGACATCAT 2530

Qy   2528 CCCGGGGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGGGCCAGGA 2587
          || | ||||||||||| |||||||||||||||||||||||||| ||  |||| |||||||
Db   2531 CCAGAGGCCACCACTCAGACCGAGAGCCACTGCTGGGGCCCCCGGCCACCCGTGCCAGGA 2590

Qy   2588 GCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGACATGGGGG 2647
          | || |||||| |||||  ||||||||||||||||| ||||||||| || ||||||||||
Db   2591 GTCTTCTCCCTTTGCCCCTGGAGGGCAAGGCCAAACGAAGCCACTCTTTTGACATGGGGG 2650

Qy   2648 ACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCTCGGAAGG 2707
```

```
               ||||||| ||||| || ||      | || || || ||||||||||| ||| ||
Db        2651 ACTTTGCAGCTGCAGCTGC------AGCTGTTCCTGGTGGCTACAGTCCCCCTAGGCGGG 2704

Qy        2708 TCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCCTTTGAGG 2767
               |||| ||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        2705 TCTCAAACATCTGGACGAAGCGCAGCCTGTCTGTCAATGGCATGCTCTTGCCCTTTGAGG 2764

Qy        2768 AGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATGGAGAGCA 2827
               |||||||||||||||| ||||||||||| ||||||||||| || ||||| ||||| || |
Db        2765 AGAGTGACCTGGTGGGAGCCCGGGGGACATTTGGCAGCTCAGAGTGGGTAATGGAAAGTA 2824

Qy        2828 CGGTCTAGGTGGGGGTGG-GCATGCTCCCTTTCCTGTGCGCAGGGTGGAGAAGGGGAAA 2886
               | || |||   |||  || | |  ||||||  |     |  || | | ||| | ||
Db        2825 CTGTGTAGCCAGGGAAGGAGTACCCTCCCTCCCACCCTCAAAGTGGGAGAGGAAGGAGTC 2884

Qy        2887 GAA--TCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGACAG 2944
               | |     ||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        2885 GGAGTCAGCACTGGCAAGTATTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGACAG 2944

Qy        2945 TTTCGTCTCCCTGTCAATGGCCTCGTGTC------CCCCCCTACCCCGCAACACCCACAT 2998
               | | |||||||||||||  ||||| ||||      |||  |||| || |    ||||  |
Db        2945 TCTTGTCTCCCTGTCAACAGCCTCATGTCCTTCTGCCCAGCTACACCCCCTTCCCCATGT 3004

Qy        2999 CACCTCCCCA-----------CCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCT 3047
               ||||| ||||            |  ||   |||||||||||||||||||| |||||||||
Db        3005 CACCTTCCCAGGCCTGTCCCCACCACCCAGCGGGGTGTGCTCAGGGAATG-GGACTCGCT 3063

Qy        3048 CAAATGCCGGACTGAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAA 3107
               ||  || | |||||||||||||||||||||| |||| |    | ||||||||| ||||||
Db        3064 CATGTGTCAGACTGAGCCCTGAGTGTTTGGAGAGGCCAACAACTGCCTTTCTATTCACAA 3123

Qy        3108 ATGTAGCCTACAAGCAAGCGGCTTTGGATTGCTTATG 3144
               |||||| | ||||||||||||||| ||||||||||||
Db        3124 ATGTAG-CGACAAGCAAGCGGCTTCGGATTGCTTATG 3159
```

```
RESULT 2
BM787163
LOCUS       BM787163                 611 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0066050 S11SNU1 Homo sapiens cDNA clone S11SNU1-20-H05 5',
            mRNA sequence.
ACCESSION   BM787163
VERSION     BM787163.1  GI:19135395
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 611)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 20  row: H  column: 05
            High quality sequence stop: 611.
FEATURES             Location/Qualifiers
     source          1..611
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="S11SNU1-20-H05"
                     /clone_lib="S11SNU1"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Lymphoblast-like"
                     /cell_line="SNU-1"
                     /lab_host="Top10F'"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT      108 a    181 c    184 g    138 t
ORIGIN

  Query Match             19.4%;  Score 609.4;  DB 14;  Length 611;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-119;
  Matches  610;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          81 ACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGTTCTTCGATCCGGAAAATCCT 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGTTCTTCGATCCGGAAAATCCT 60

Qy         141 ACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTTCTTTAATCTGGAAGAGAAGA 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTTCTTTAATCTGGAAGAGAAGA 120

Qy         201 GAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGCCATGGATATAACTGAATAAG 260
               ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         121 GAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAATGCCATGGATATAACTGAATAAG 180

Qy         261 CGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCATCTGCCTGCATTCGCCGGAG 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCATCTGCCTGCATTCGCCGGAG 240

Qy         321 CCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGG 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGG 300

Qy         381 GAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCAGCCGGGGGTAGATGCTGCCT 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCAGCCGGGGGTAGATGCTGCCT 360

Qy         441 CGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGGCCTGCTAGCGTTTGG 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGGCCTGCTAGCGTTTGG 420

Qy         501 CATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGA 560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CATGGCGTTTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGA 480

Qy         561 GTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCG 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GTCACTGGGGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCG 540

Qy         621 GCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTT 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GCGGACAGTGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTT 600

Qy         681 TGCCAACATGA 691
               |||||||||||
Db         601 TGCCAACATGA 611

RESULT 3
BF312768
LOCUS       BF312768                 621 bp    mRNA    linear   EST 21-NOV-2000
DEFINITION  601896906F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4126224 5',
            mRNA sequence.
```

```
ACCESSION   BF312768
VERSION     BF312768.1  GI:11260634
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM1015  row: 1  column: 01
            High quality sequence stop: 621.
FEATURES             Location/Qualifiers
     source          1..621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4126224"
                     /clone_lib="NIH_MGC_19"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      118 a    207 c    168 g    128 t
ORIGIN

  Query Match             19.3%;  Score 606.8;  DB 12;  Length 621;
  Best Local Similarity   99.5%;  Pred. No. 4.3e-119;
  Matches  619;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         806 GCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACA 865
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GCCTTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACA 60

Qy         866 ACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATC 925
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ACCAGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATC 120

Qy         926 TGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCA 985
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TGGACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCA 180

Qy         986 ACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTG 1045
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ACCTCCACCAGCTGAGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTG 240

Qy        1046 CAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCC 1105
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CAGACCTGCAGAAACTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCC 300

Qy        1106 CTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCAC 1165
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CTGATCCCATCTTTGCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCAC 360

Qy        1166 CCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGC 1225
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CCTTGTCCTTTAGTTTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGC 420

Qy        1226 GGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTC 1285
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTC 480

Qy        1286 GCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGC 1345
               ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         481 GCTACTTCTGGCATGTGCGTGAGGAGGAGTTTG-GTGCGAGCCGCCTCTCATCACCCAGC 539

Qy        1346 ACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTG 1405
               |||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db         540 ACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGTGCACACTCAAGTGCAAAGCCATTG 599

Qy        1406 GGGACCCCAGCCCCCTTATCCA 1427
               ||||||||||||||||||||||
Db         600 GGGACCCCAGCCCCCTTATCCA 621


RESULT 4
BQ126354
LOCUS       BQ126354                 582 bp    mRNA    linear   EST 19-APR-2002
DEFINITION  ii16e11.y1 Melton Amplified Mouse E16 5 Pancreas 3 M16S1 A Mus
            musculus cDNA clone IMAGE:5941749 5' similar to TR:Q9ULH4 Q9ULH4
            KIAA1246 PROTEIN ;, mRNA sequence.
ACCESSION   BQ126354
VERSION     BQ126354.1  GI:20200265
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ii16e11.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center This clone is
            available royalty-free through LLNL; please contact the IMAGE
            consortium (info@image.llnl.gov) for further information
            Seq primer: -40RP from Gibco
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1..582
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5941749"
                     /clone_lib="Melton Amplified Mouse E16 5 Pancreas 3 M16S1
                     A"
                     /sex="Both"
                     /tissue_type="pancreas"
                     /dev_stage="Embryonic day E16.5"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1;
                     Library constructed using SuperScript Plasmid Library kit
                     (Life Technologies). cDNA made by oligo-dT priming.
                     Size-selected by column fractionation; average insert size
                     0.97 kb. Amplified once on solid support. cDNA Library
                     Preparation: Guolin Chen."
BASE COUNT      126 a    194 c    158 g    104 t
ORIGIN
```

```
  Query Match             18.5%;  Score 582;  DB 14;  Length 582;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-114;
  Matches  582;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1408 GACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCA 1467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACCCCAGCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCA 60

Qy       1468 AGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGT 1527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGT 120

Qy       1528 GGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTC 1587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTC 180

Qy       1588 TCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCC 1647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCC 240

Qy       1648 CGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGC 1707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGC 300

Qy       1708 GGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACC 1767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACC 360

Qy       1768 TCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTG 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTG 420

Qy       1828 CAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAG 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATCCCAGCCTCCAACAAG 480

Qy       1888 GCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATG 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTTCGTGGTCAACAACCTGGTGTCAGGGACTGGCTACGACTTGTGTGTGCTGGCCATG 540

Qy       1948 TGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTG 1989
              ||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGGATGACACAGCCACGACACTCACGGCCACCAACATCGTG 582


RESULT 5
BQ126066/c
LOCUS       BQ126066                 578 bp    mRNA    linear   EST 19-APR-2002
DEFINITION  ii16e11.x1 Melton Amplified Mouse E16 5 Pancreas 3 M16S1 A Mus
            musculus cDNA clone IMAGE:5941749 3' similar to TR:Q9ULH4 Q9ULH4
            KIAA1246 PROTEIN ;, mRNA sequence.
ACCESSION   BQ126066
VERSION     BQ126066.1  GI:20199977
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center This clone is
            available royalty-free through LLNL; please contact the IMAGE
            consortium (info@image.llnl.gov) for further information
            Seq primer: -40UP from Gibco
            High quality sequence stop: 388.
FEATURES             Location/Qualifiers
     source          1..578
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5941749"
                     /clone_lib="Melton Amplified Mouse E16 5 Pancreas 3 M16S1
                     A"
                     /sex="Both"
                     /tissue_type="pancreas"
                     /dev_stage="Embryonic day E16.5"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1;
                     Library constructed using SuperScript Plasmid Library kit
                     (Life Technologies). cDNA made by oligo-dT priming.
                     Size-selected by column fractionation; average insert size
                     0.97 kb. Amplified once on solid support. cDNA Library
                     Preparation: Guolin Chen."
BASE COUNT      121 a    182 c    161 g    114 t
ORIGIN

  Query Match             17.2%;  Score 540.2;  DB 14;  Length 578;
  Best Local Similarity   98.4%;  Pred. No. 6.9e-105;
  Matches  556;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy       2575 GCCCGGGCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCC 2634
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        578 GCCCGGGCCAGGAGCCTGCTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCACCCACTCC 519

Qy       2635 TTCGACAT-GGGGGACTTTGCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAG 2693
              |||||||| ||||||||||| |||||||||||||||||||||||| |||||||||| |||
Db        518 TTCGACATGGGGGGACTTTGTTGCTGCGGCGGCGGGAGGGGTCGTTCCGGGCGGCTCCAG 459

Qy       2694 TCCTCCTCGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCT 2753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 TCCTCCTCGGAAGGTCTCGAACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCT 399

Qy       2754 CTTGCCCTTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATG 2813
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 CTTGCCCTTTGAGGAGAGTGACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATG 339

Qy       2814 GGTGATGGAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTG 2873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 GGTGATGGAGAGCACGGTCTAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTG 279

Qy       2874 GGAGAAGGGGAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACA 2933
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GGAGAAGGGGAAAGAATCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACA 219

Qy       2934 TCCAGGGACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACC 2993
              || |   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 TCAAAAAACAGTTTCGTCTCCCTGTCAATGGCCTCGTGTCCCCCCCTACCCCGCAACACC 159

Qy       2994 CACATCACCTCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATG 3053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 CACATCACCTCCCCACCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAATG 99

Qy       3054 CCGGACTGAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAG 3113
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 CCGGACTGAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTAG 39

Qy       3114 CCTACAAGCAAGCGGCTTTGGATTG 3138
              |||||||||||||||||||||||||
Db         38 CCTACAAGCAAGCGGCTTTGGATTG 14
```

```
RESULT 6
BM948451
LOCUS       BM948451                 652 bp    mRNA    linear   EST 14-MAR-2002
DEFINITION  UI-M-EG0p-bvd-m-12-0-UI.r1 NIH_BMAP_EG0p Mus musculus cDNA clone
            IMAGE:5690675 5', mRNA sequence.
ACCESSION   BM948451
VERSION     BM948451.1  GI:19432042
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, Univeristy of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1..652
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5690675"
                     /clone_lib="NIH_BMAP_EG0p"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 18.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured mRNa was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was size selected according to mRNA size fraction,
                     ligated with EcoR I adaptor, digested with Not I, and then
                     cloned directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     , is CAGCCACGAC. This library was created for the
                     University of Iowa Mouse Brain Molecular Anatomy Project
                     (BMAP): 'Gene Discovery in the Developing Mouse Nervous
                     System', supported by National Institute of Mental Health
                     (NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT      143 a    183 c    189 g    137 t
ORIGIN

  Query Match             16.7%;  Score 525.6;  DB 14;  Length 652;
  Best Local Similarity   87.9%;  Pred. No. 9.2e-102;
  Matches  573;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy       1221 GCTGCGGAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAA 1280
              |||||||||||| ||| |||| ||||||||||| |||||||| ||||| ||  | |||||
Db          1 GCTGCGGAGGCTGGAGAGGGATGATGACCTGGAGACCTGTGGATCCCCCGGAAGTCTCAA 60

Qy       1281 GGGTCGCTACTTCTGGCATGTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCAC 1340
              |||||||||||| |||||| | |||||||||||||||||||| |||||||||||||||||
Db         61 GGGTCGCTACTTTTGGCATATTCGTGAGGAGGAGTTTGTGTGTGAGCCGCCTCTCATCAC 120

Qy       1341 CCAGCACACACACAAGTTGCTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGC 1400
              |||||||||||||||| |||||||||||||||||||||| ||||| || |||||||| ||
Db        121 CCAGCACACACACAAGCTGCTGGTTCTGGAGGGCCAGGCAGCCACTCTTAAGTGCAAGGC 180

Qy       1401 CATTGGGACCCCAGCCCCCCTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAA 1460
              ||||||||||||||||||  | ||||||||||| ||||||||||||||| |||| || ||
Db        181 CATTGGGACCCCAGCCCCTTTGATCCACTGGGTTGCCCCCGATGACCGCTTGGTGGGAAA 240

Qy       1461 CTCCTCAAGGACCGCTGTCTATGACAATGGCACCCTGGACATCTTCATCACCACATCTCA 1520
              ||| || ||||| ||||| |||||||||||||||||||||||  | || ||||| |||||
Db        241 CTCTTCCAGGACTGCTGTATATGACAATGGCACCCTGGACATTCTAATTACCACCTCTCA 300

Qy       1521 GGACAGTGGTGCCTTCACCTGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGT 1580
              |||||||||  |||| |||||||| || |||||||| || |||||||| |||||||  ||
Db        301 GGACAGTGGACCCTTTACCTGCATCGCAGCCAATGCAGCAGGAGAGGCTACGGCCACTGT 360

Qy       1581 GGAGGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCC 1640
               ||||||||||| || ||||| |||||||| || |||||||| ||||| |  ||||||||
Db        361 AGAGGTCTCCATTGTGCAGCTCCCACACCTTAGTAACAGCACTAGCCGGATGGCACCCCC 420

Qy       1641 CAAGTCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAG 1700
              |||||| ||||| || |||||||| || ||||||||||||||||||||||| ||||| ||
Db        421 CAAGTCTCGCCTTTCGGACATCACAGGTTCCAGCAAGACCAGCCGGGGAGGGGGAGGTAG 480

Qy       1701 TGGGGGCGGAGAGCCTCCCAAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGAC 1760
              |||||  || ||||||||||||||| |||| ||  |||||||||||||||| || || ||
Db        481 TGGGGCTGGGGAGCCTCCCAAAAGCACCCCAGAGAGGGCTGTGCTTGTGTCCGATGTTAC 540

Qy       1761 CACCACCTCGGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTA 1820
              |||||| || |||||||||||||||||||||||||||||||||||| |||||||||||||
Db        541 CACCACATCTGCCCTGGTCAAGTGGTCTGTCAGCAAGTCAGCACCCAGGGTGAAGATGTA 600

Qy       1821 CCAGCTGCAGTACAACTGCTCTGACGATGAGGTACTGATTTACAGGATGATC 1872
               ||||||||||||||||| ||||| |||||||| ||||| ||||||||||||
Db        601 TCAGCTGCAGTACAACTGTTCTGATGATGAGGTGCTGATCTACAGGATGATC 652


RESULT 7
BB642310
LOCUS       BB642310                 655 bp    mRNA    linear   EST 31-AUG-2001
DEFINITION  BB642310 RIKEN full-length enriched, adult retina Mus musculus cDNA
            clone A930009O11 5', mRNA sequence.
ACCESSION   BB642310
VERSION     BB642310.1  GI:15401853
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
```

```
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T.,
            Ishii,Y. and Hayashizaki,Y.
             Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
            Func. Genomics 2 pre, L72-L86 (2001
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A930009O11"
                     /clone_lib="RIKEN full-length enriched, adult retina"
                     /tissue_type="retina"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5'GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. ~Retina RNA was provided by Stefano Gustincich,
                     Department of Neurobiology, Harvard Medical School, 220
                     Longwood Ave., Boston, MA02115, USA, whose assistance we
                     gratefully acknowledge."
BASE COUNT      119 a    203 c    188 g    144 t      1 others
ORIGIN

  Query Match             16.4%;  Score 514.2;  DB 10;  Length 655;
  Best Local Similarity   87.8%;  Pred. No. 2.5e-99;
  Matches  574;  Conservative    0;  Mismatches   74;  Indels    6;  Gaps    1;

Qy        335 AGCTCGAGTCTGTCTCGGGCGGGGAAGGATGCGTGGCCGAGCCGGGGAGCCCGGGCGCCC 394
              |||  ||| ||||||| |||  |||||||||||||||||||||  |||||||||||||||
Db          2 AGCGTGAGCCTGTCTCCGGCCAGGAAGGATGCGTGGCCGAGCCAAGGAGCCCGGGCGCCC 61

Qy        395 CGCGGAGCCGGCCTCGGTGCCACCCAGCCGGGGGTAGATGCTGCCTCGCCCAGG------ 448
               ||||||||||||| || ||| ||| |||  |||||||||| |||| |||||||
Db         62 TGCGGAGCCGGCCTTGGAGCCTCCCCGCCACGGGTAGATGCCGCCTTGCCCAGGTGCAGA 121

Qy        449 CGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGGCCTGCTAGCGTTTGGCATGGCGT 508
               |||||||||||||||||||||||| ||||||||||| ||||| || |||||||||||||
Db        122 GGCTGAGTGACCAGACCATGGAGACTCTGCTTGGTGGGCTGCTGGCTTTTGGCATGGCGT 181

Qy        509 TTGCCGTGGTCGACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGG 568
              |||| |||||||| ||||||||||||||||| ||||||||||||||||||||||||||||
Db        182 TTGCTGTGGTCGATGCCTGCCCCAAGTACTGCGTCTGCCAGAATCTGTCTGAGTCACTGG 241

Qy        569 GGACCCTGTGCCCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAG 628
              |||||||||| |||||||||||||| |||||||| |||||||| ||||||||| |||| |
Db        242 GGACCCTGTGTCCCTCCAAGGGGCTCCTCTTTGTGCCCCCTGACATTGACCGGAGGACGG 301

Qy        629 TGGAGCTGCGCCTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACA 688
              |||||||||| ||||| ||||||||||||||||||||  ||||||| |||||||||||||
Db        302 TGGAGCTGCGTCTGGGTGGCAACTTCATCATCCACATTGGCCGCCAAGACTTTGCCAACA 361

Qy        689 TGACGGGGCTGGTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTT 748
              |||| ||||||||||| |||||| |||||||||||||||||||||| ||||||||||| |
Db        362 TGACAGGGCTGGTGGATCTGACCTTGTCCAGGAACACCATCAGCCATATCCAGCCCTTCT 421

Qy        749 CCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCC 808
              |||| |||||||| |||||||||||||||||||| ||||||||||| ||||| || ||||
Db        422 CCTTCCTGGACCTTGAGAGCCTCCGCTCCCTGCACCTTGACAGCAACCGGCTACCTAGCC 481

Qy        809 TTGGGGAGGACACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACC 868
              ||||||||||||| |||||||| |||||||||||||| |||||||| |||||||| ||||
Db        482 TTGGGGAGGACACACTCCGGGGTCTGGTCAACCTGCAACACCTTATTGTGAACAATAACC 541

Qy        869 AGCTGGGCGGCATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGG 928
              ||||||| |||||||| ||||| || ||||| |||||| ||||||| || || || ||||
Db        542 AGCTGGGTGGCATCGCTGATGATGCCTTTGAAGACTTCTTGCTGACTTTAGAAGACCTGG 601

Qy        929 ACCTCTCCTACAACAACCTCCATGGCCTGCCGTGGGACTCCGTGCGACGCATGG 982
              |||| || |||||  ||||||| || || || || || ||||| || |||||||
Db        602 ACCTATCTTACAAACACCTCCACGGACTTCCCTGNGATTCCGTACGGCGCATGG 655


RESULT 8
BC007718
LOCUS       BC007718                1853 bp    mRNA    linear   HTC 12-JUL-2001
DEFINITION  Homo sapiens, hypothetical protein MGC3103, clone IMAGE:4110922,
            mRNA.
ACCESSION   BC007718
VERSION     BC007718.1  GI:14712783
KEYWORDS    HTC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1853)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site:         http://www.nisc.nih.gov/
            Contact:          nisc_mgc@nhgri.nih.gov
            Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
            Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 17 Row: e Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: GenomeScan gene prediction
            This clone has the following problem: incomplete processing.
FEATURES             Location/Qualifiers
     source          1. .1853
                     /organism="Homo sapiens"
                     /db_xref="LocusID:78999"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4110922"
```

```
                  /tissue_type="Muscle, rhabdomyosarcoma"
                  /clone_lib="NIH_MGC_17"
                  /lab_host="DH10B-R"
                  /note="Vector: pOTB7"
BASE COUNT      329 a    648 c    558 g    318 t
ORIGIN

  Query Match            15.7%;  Score 492.6;  DB 11;  Length 1853;
  Best Local Similarity  63.2%;  Pred. No. 1.1e-94;
  Matches  828;  Conservative    0;  Mismatches  444;  Indels   39;  Gaps    3;

Qy        520 GACGCCTGCCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACTGGGGACCCTGTGC 579
              | |||||||||   |   || |||||||||||| ||||| ||||| ||  | ||||| ||
Db        272 GCCGCCTGCCCGCTGCCCTGCGTCTGCCAGAACCTGTCCGAGTCGCTCAGCACCCTCTGT 331

Qy        580 CCCTCCAAGGGGCTGCTCTTTGTACCCCCTGATATTGACCGGCGGACAGTGGAGCTGCGC 639
               || |    || ||||| |||| || ||  |  | |||||||| |||||||||||||||
Db        332 GCCCACCGAGGCCTGCTGTTTGTGCCGCCCAACGTGGACCGGCGCACAGTGGAGCTGCGG 391

Qy        640 CTGGGCGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTG 699
              ||||  | ||||||||||     | |  | | ||  |||||   |||||||||||| |||
Db        392 CTGGCTGACAACTTCATCCAGGCCCTGGGGCCCCCTGACTTCCGCAACATGACGGGACTG 451

Qy        700 GTGGACCTGACCCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGAC 759
              ||||||||||| |||||  | ||  |||||| || |||   | ||    |||||  ||||
Db        452 GTGGACCTGACACTGTCTCGCAATGCCATCACCCGCATTGGGGCCCGCGCCTTTGGGGAC 511

Qy        760 CTCGAGAGCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGAC 819
              ||||||||||| || ||||| || |||||| ||||  |||||      || ||    |
Db        512 CTCGAGAGCCTGCGTTCCCTCCACCTTGACGGCAACAGGCTGGTGGAGCTGGGCACCGGG 571

Qy        820 ACCCTCCGGGGCCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGC 879
              | |||||||||||  ||||| |||||||||||| ||| | | |  |||||||||||||| ||
Db        572 AGCCTCCGGGGCCCCGTCAATCTGCAGCACCTCATCCTCAGCGGCAACCAGCTGGGCCGC 631

Qy        880 ATCGCAGATGAGGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTAC 939
              |||||    |  || || || |||||||||   ||   ||||||| |||||||||| ||||||
Db        632 ATCGCGCCGGGAGCCTTCGACGACTTCCTAGAGAGCCTGGAGGACCTGGACCTGTCCTAC 691

Qy        940 AACAACCTCCATGGCCTGCCGTGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTG 999
              ||||||||||      |||| |||| |  | |  |   ||||     ||| |||   ||
Db        692 AACAACCTCCGGCAGGTGCCCTGGCCGGCATCGGCGCCATGCCTGCCCTGCACACCCTC 751

Qy       1000 AGCCTGGACCACAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAA 1059
              | ||||||||| |||||  | ||     |  ||   ||| |||| ||  | ||     |
Db        752 AACCTGGACCATAACCTTATTGACGCACTGCCCCCAGGCGCCTTCGCCCAGCTCGGTCAG 811

Qy       1060 CTGGCCCGCCTGGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTT 1119
              ||  |||||||||| ||||||||||| || |||   | ||||| ||  || || ||  | ||
Db        812 CTCTCCCGCCTGGACCTCACCTCCAACCGCCTGGCCACGCTGGCTCCGGACCCGCTTTTC 871

Qy       1120 GCCCGCTCCCAGGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGT 1179
               | ||    |  | |            ||     ||| | || | ||| ||    | ||
Db        872 TCTCGTGGGCGTGAT------------GCAGAGGCCTCTCCCGCCCCCCTGGTGCTGAGC 919

Qy       1180 TTTGGGGGTAACCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGG 1239
              ||| | || ||||| || ||||||||| |||||||| || |||||||||| |||| | ||||
Db        920 TTTAGCGGGAACCCCCTGCACTGCAACTGTGAGCTGCTGTGGCTGCGGCGGCTGGCGCGG 979

Qy       1240 GACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCAT 1299
                 || ||||||||||| || | ||||||    |||||    || ||||||||||||
Db        980 CCGGACGACCTGGAAACGTGCGCCTCCCCGCCCGGCCTGGCCGGCCGCTACTTCTGGGCA 1039

Qy       1300 GTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTG 1359
              ||||  ||||  |||||    || ||||||||| |||||  |||  |||||  ||     |
Db       1040 GTGCCCGAGGGCGAGTTCTCCTGTGAGCCGCCCCTCATTGCCCGCCACACGCAGCGCCTC 1099

Qy       1360 CTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCC 1419
                ||| ||||| ||||||  |||||| ||   ||||   ||| | || ||||||| | ||
Db       1100 TGGGTGCTGGAAGGCCAGCGGGCCACGCTGCGGTGCCGGGCCCTGGGTGACCCCGCGCCT 1159

Qy       1420 CTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTC 1479
                 || |||||||| |  || || |||||  |||| || ||||||||  |  ||   |
Db       1160 ACCATGCACTGGGTCGGTCCTGACGACCGGTTGGTTGGCAACTCCTCCCGAGCCCGGGCT 1219

Qy       1480 TATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 1539
              |    ||| || ||| | ||  ||     | |||    ||  ||||  ||| | || ||||
Db       1220 TTCCCCAACGGGACCTTAGAGATTGGGGTGACCGGCGCTGGGGACGCTGGGGGCTACACC 1279

Qy       1540 TGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCATCGTCCAG 1599
              ||||| ||  ||||  |||| || |||||||| || ||    || ||  |     |  |
Db       1280 TGCATCGCCACCAACCCTGCTGGTGAGGCCACAGCCCGAGTAGAACTGCGGGTGCTGGCC 1339

Qy       1600 CTGCCACACCTCAGCAACAGCACCAGCCGCACTGCACCCCCCAAGTCCCGCCTCTCAGAC 1659
               |||| ||     | |||||||    |      |  | ||      || | | ||| |||
Db       1340 TTGCCCATGGTGGGAACAGCAGTGCCGAGGGGGGCCGCC------CCGGGCCCTCGGAC 1393

Qy       1660 ATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGGAGGCAGTGGGGGCGGAGAGCCTCCC 1719
              ||| | | ||||      ||       |  ||||  || |    ||     |||||
Db       1394 ATCGCCGCCTCCGCTCGCACTGCTGCCGAGGGTGAGGGGACGCTGGAGTCTGAGCC---- 1449

Qy       1720 AAAAGCCCCCCGGAACGGGCTGTGCTTGTGTCTGAAGTGACCACCACCTCGGCCCTGGTC 1779
                                ||  ||||  ||| | || |||||| ||||||| |  |||||
Db       1450 -----------------AGCCGTGCAGGTGACGGAGGTGACCGCCACCTCAGGGCTGGTG 1492

Qy       1780 AAGTGGTCTGTCAGCAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAG 1830
              |  |||  |  | |   | ||| |  ||| |||  ||||| |||  | |||
Db       1493 AGCTGGGGTCCCGGGCGGCCAGCCGACCCAGTGTGGATGTTCCAAATCCAG 1543

RESULT 9
AW163387
LOCUS       AW163387                 457 bp    mRNA    linear   EST 09-NOV-1999
DEFINITION  au94f03.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone
            IMAGE:2783933 5', mRNA sequence.
ACCESSION   AW163387
VERSION     AW163387.1  GI:6302420
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: au94f03.x1
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 435.
FEATURES             Location/Qualifiers
     source          1..457
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2783933"
                     /clone_lib="Schneider fetal brain 00004"
                     /sex="male"
                     /tissue_type="frontal lobe"
                     /dev_stage="5 months post-conception"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescript SK (Stratagene);
                     Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
                     prepared from human fetal brain tissue.  5' and 3'
                     adaptors were used in cloning as follows: 5' adaptor
                     sequence:
                     5'-GAGAGAGAGAAGAGCTCAAGGATCCTTAATTAAATTAATCCCCCCCCCCCC-3'
```

```
                      and 3' adaptor sequence:
                      5'-GAGAGAGAGACTCGAGTTTTTTTTTTTTTTTTTT-3'.  The library was
                      size-selected for >0.5 kb inserts and has an average
                      insert size estimated at 1.2 kb.  This library was
                      constructed using the CAP-trapper method for full-length
                      enrichment and has not undergone amplification.    Library
                      was constructed by Dr. Claudio Schneider (LNCIB-Area
                      Science Park,  Trieste, Italy). "
BASE COUNT       83 a    131 c    140 g    103 t
ORIGIN

  Query Match             14.1%;  Score 441.8;  DB 10;  Length 457;
  Best Local Similarity   99.3%;  Pred. No. 6.7e-84;
  Matches  454;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         66 ACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGTTCTTC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGTTCTTC 60

Qy        126 GATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTTCTTTA 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTTCTTTA 120

Qy        186 ATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGCCATGG 245
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        121 ATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAATGCCATGG 180

Qy        246 ATATAACTGAAT-AAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCATCTG 304
              |||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 ATATAACTGAATAAAGCCGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCATCTG 240

Qy        305 CCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAAGGAT 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAAGGAT 300

Qy        365 GCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCAGCCG 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCAGCCG 360

Qy        425 GGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTG 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTG 420

Qy        485 GCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGA 521
              |||||||||||||||||||||||||||||||||||||
Db        421 GCCTGCTAGCGTTTGGCATGGCGTTTGCCGTGGTCGA 457


RESULT 10
AW161566
LOCUS       AW161566                 435 bp    mRNA    linear   EST 09-NOV-1999
DEFINITION  au71e01.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone
            IMAGE:2781720 5', mRNA sequence.
ACCESSION   AW161566
VERSION     AW161566.1  GI:6300599
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 435)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: au71e01.x1
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1..435
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2781720"
                     /clone_lib="Schneider fetal brain 00004"
                     /sex="male"
                     /tissue_type="frontal lobe"
                     /dev_stage="5 months post-conception"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescript SK (Stratagene);
                     Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
                     prepared from human fetal brain tissue.  5' and 3'
                     adaptors were used in cloning as follows: 5' adaptor
                     sequence:
                     5'-GAGAGAGAGAGCTCAAGGATCCTTAATTAAATTAATCCCCCCCCCCCC-3'
                     and 3' adaptor sequence:
                     5'-GAGAGAGAGACTCGAGTTTTTTTTTTTTTTTTTT-3'.  The library was
                     size-selected for >0.5 kb inserts and has an average
                     insert size estimated at 1.2 kb.  This library was
                     constructed using the CAP-trapper method for full-length
                     enrichment and has not undergone amplification.    Library
                     was constructed by Dr. Claudio Schneider (LNCIB-Area
                     Science Park,  Trieste, Italy). "
BASE COUNT       80 a    126 c    131 g     97 t      1 others
ORIGIN

  Query Match             13.4%;  Score 422.8;  DB 10;  Length 435;
  Best Local Similarity   98.2%;  Pred. No. 7.5e-80;
  Matches  427;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         66 ACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGTTCTTC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACTTGAAGACACCCCACATTCCAAGATGCCCGAGGTTCCTGGGAATGCCTGGGGTTCTTC 60

Qy        126 GATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTTCTTTA 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATCCGGAAAATCCTACCGGCATCCTCCTAGGGAGGGATTATTATTATTATTTTTCTTTA 120

Qy        186 ATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAACGCCATGG 245
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        121 ATCTGGAAGAGAAGAGAACAAGTTGTGCTTTTCCCCCCTTCTTCTTGCTAAATGCCATGG 180

Qy        246 ATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCATCTGC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATATAACTGAATAAGCGGCTCAGGGCTTTCCCCGCGTGGACGTCCGAGGCCACCATCTGC 240

Qy        306 CTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAAGGATG 365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCATTCGCCGGAGCCGCCGGAGGGTTTAGCTCGAGTCTGTCTCGGGCGGGGAAGGATG 300

Qy        366 CGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCAGCCGG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   | |
Db        301 CGTGGCCGAGCCGGGGAGCCCGGGCGCCCCGCGGAGCCGGCCTCGGTGCCACCCNAGCCG 360

Qy        426 GGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGG 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGTAGATGCTGCCTCGCCCAGGCGCTGAGTGACCAGACCATGGAGACCCTGCTTGGTGG 420

Qy        486 CCTGCTAGCGTTTGG 500
              |||||   |||||||
Db        421 CCTGCCTACGTTTGG 435


RESULT 11
BB278209
LOCUS       BB278209                 685 bp    mRNA    linear   EST 31-AUG-2001
DEFINITION  BB278209 RIKEN full-length enriched, adult retina Mus musculus cDNA
```

```
            clone A930009O11 3', mRNA sequence.
ACCESSION   BB278209
VERSION     BB278209.2  GI:15410982
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Jul 7, 2000 this sequence version replaced gi:8975231.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T.,
            Ishii,Y. and Hayashizaki,Y.
             Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
            Func. Genomics 2 pre, L72-L86 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1..685
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A930009O11"
                     /clone_lib="RIKEN full-length enriched, adult retina"
                     /tissue_type="retina"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5'GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. -Retina RNA was provided by Stefano Gustincich,
                     Department of Neurobiology, Harvard Medical School, 220
                     Longwood Ave., Boston, MA02115, USA, whose assistance we
                     gratefully acknowledge."
BASE COUNT      146 a    199 c    196 g    144 t
ORIGIN

  Query Match             12.0%;  Score 377.2;  DB 10;  Length 685;
  Best Local Similarity   78.9%;  Pred. No. 4.3e-70;
  Matches  546;  Conservative    0;  Mismatches  118;  Indels   28;  Gaps    7;

Qy       2473 AAGGAGGAGCTGCTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGGGACGTCGGCCCGG 2532
              || || || || || ||||||||||||||||| || |||||||| ||||| ||  || | 
Db          1 AAAGAAGAACTTCTAGACTCCAGGACTCCAGCAGGCAGAGGGGCAGGGACATCATCCAGA 60

Qy       2533 GGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGGGCCAGGAGCCTG 2592
              ||||||||||| |||||||||||||||||||||||||| ||  |||| |||||||| || 
Db         61 GGCCACCACTCAGACCGAGAGCCACTGCTGGGGCCCCCGGCCACCCGTGCCAGGAGTCTT 120

Qy       2593 CTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGACATGGGGGACTTT 2652
              ||||| |||||  ||||||||||||||||||| |||||||| || |||||||||||||||
Db        121 CTCCCTTTGCCCCTGGAGGGCAAGGCCAAACGAAGCCACTCTTTTGACATGGGGGACTTT 180

Qy       2653 GCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCTCGGAAGGTCTCG 2712
              || ||||| || ||       |  || || || ||||||||||| ||| ||  |||||| 
Db        181 GCAGCTGCAGCTGC------AGCTGTTCCTGGTGGCTACAGTCCCCCTAGGCGGGTCTCA 234

Qy       2713 AACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCCTTTGAGGAGAGT 2772
              ||||||||||||||||||||||| |||||||| |||||||||||||||||||||||||||
Db        235 AACATCTGGACGAAGCGCAGCCTGTCTGTCAATGGCATGCTCTTGCCCTTTGAGGAGAGT 294

Qy       2773 GACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATGGAGAGCACGGTC 2832
              ||||||||||| ||||||||||| ||||||||||| || ||||| ||||| || || || 
Db        295 GACCTGGTGGGAGCCCGGGGGACATTTGGCAGCTCAGAGTGGGTAATGGAAAGTACTGTG 354

Qy       2833 TAGGTGGGGGTGG-GCATGCTCCCTTTCCTGTGCGCAGGGTGGGAGAAGGGGAAAGAA-- 2889
              |||   |||  || | |  ||||||  |     |  || | | ||| | ||    | |  
Db        355 TAGCCAGGGAAGGAGTACCCTCCCTCCCACCCTCAAAGTGGGAGAGGAAGGAGTCGGAGT 414

Qy       2890 TCTCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGACAGTTTCG 2949
                 ||||||||||| ||||||||||||||||||||||||||||||||||||||||| | |
Db        415 CAGCACTGGCAAGTATTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGACAGTCTTG 474

Qy       2950 TCTCCCTGTCAATGGCCTCGTGTC------CCCCCCTACCCCGCAACACCCACATCACCT 3003
              ||||||||||||  ||||| ||||      |||  |||| || |    ||||  ||||||
Db        475 TCTCCCTGTCAACAGCCTCATGTCCTTCTGCCCAGCTACACCCCCTTCCCCATGTCACCT 534

Qy       3004 CCCCA-----------CCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTCAAAT 3052
               ||||            |  ||   |||||||||||||||||||| |||||||||||  |
Db        535 TCCCAGGCCTGTCCCCACCACCCAGCGGGGTGTGCTCAGGGAATG-GGACTCGCTCATGT 593

Qy       3053 GCCGGACTGAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAATGTA 3112
              | | |||||||||||||||||||||| |||| |    | ||||||||| |||||||||||
Db        594 GTCAGACTGAGCCCTGAGTGTTTGGAGAGGCCAACAACTGCCTTTCTATTCACAAATGTA 653

Qy       3113 GCCTACAAGCAAGCGGCTTTGGATTGCTTATG 3144
              | | ||||||||||||||| ||||||||||||
Db        654 G-CGACAAGCAAGCGGCTTCGGATTGCTTATG 684


RESULT 12
BM193100
LOCUS       BM193100                 357 bp    mRNA    linear   EST 13-DEC-2001
DEFINITION  TCBAP1D9544 Pediatric pre-B cell acute lymphoblastic leukemia
```

```
            Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP9544, mRNA
            sequence.
ACCESSION   BM193100
VERSION     BM193100.1  GI:17651296
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 357)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
            Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org
            Seq primer: M13 primer.
FEATURES             Location/Qualifiers
     source          1..357
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCBAP9544"
                     /clone_lib="Pediatric pre-B cell acute lymphoblastic
                     leukemia Baylor-HGSC project=TCBA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="pre-B cell"
                     /dev_stage="pediatric 2 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
                     was primed with a BamH1-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamH1 and XhoI
                     and directionally cloned into the BamH1 and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, SasakiN, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y, High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper.,
                     DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT       71 a    111 c    103 g     72 t
ORIGIN

  Query Match             11.0%;  Score 346.6;  DB 13;  Length 357;
  Best Local Similarity   98.9%;  Pred. No. 1.3e-63;
  Matches  349;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1240 GACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCAT 1299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACGATGACCTGGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCAT 60

Qy       1300 GTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTG 1359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCGTGAGGAGGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTG 120

Qy       1360 CTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCC 1419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGTTCTGGAGGGCCAGGCGGCCACACTCAAGTGCAAAGCCATTGGGGACCCCAGCCCC 180

Qy       1420 CTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTC 1479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTATCCACTGGGTAGCCCCCGATGACCGCCTGGTAGGGAACTCCTCAAGGACCGCTGTC 240

Qy       1480 TATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 1539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TATGACAATGGCACCCTGGACATCTTCATCACCACATCTCAGGACAGTGGTGCCTTCACC 300

Qy       1540 TGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGCCATGGTGGAGGTCTCCAT 1592
              |||||||||||||||||||||||||||||||||| ||||||||||||   |||
Db        301 TGCATTGCTGCCAATGCTGCCGGAGAGGCCACGGTCATGGTGGAGGTTCTCAT 353


RESULT 13
BG380772/c
LOCUS       BG380772                 635 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  UI-R-CT0-bub-c-04-0-UI.s1 UI-R-CT0 Rattus norvegicus cDNA clone
            UI-R-CT0-bub-c-04-0-UI 3', mRNA sequence.
ACCESSION   BG380772
VERSION     BG380772.1  GI:13305244
KEYWORDS    EST.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to verify it as a clone from the
            normalized rat brain pool library cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: clones will be available through
            Research Genetics (www.resgen.com)
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..635
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-CT0-bub-c-04-0-UI"
                     /clone_lib="UI-R-CT0"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CT0
                     library is a normalized library constructed from the
                     following rat brain tissues: embryonic day 17, embryonic
                     day 19, embryonic day 21, adult day 1, adult day 12, adult
                     day 75, adult day 200. For a detailed description of the
                     library from which this clone was derived, please visit
                     our web site at ratest.eng.uiowa.edu. The subtraction has
                     been previously described in (Bonaldo, Lennon and Soares,
                     Genome Research 6:791-806, 1996)
                     TAG_LIB=UI-R-CT0
                     TAG_TISSUE=rat brain pool
                     TAG_SEQ=ACTTC"
BASE COUNT      135 a    174 c    183 g    143 t
ORIGIN

  Query Match             10.4%;  Score 327;  DB 12;  Length 635;
  Best Local Similarity   77.4%;  Pred. No. 2.2e-59;
  Matches  485;  Conservative    0;  Mismatches  110;  Indels   32;  Gaps    6;
```

```
Qy       2533 GGCCACCACTCGGACCGAGAGCCACTGCTGGGGCCCCCTGCGGCCCGGGCCAGGAGCCTG 2592
              ||||||||||| |||||||||||||||||||||||||| ||  |||| |||||||||||
Db        635 GGCCACCACTCAGACCGAGAGCCACTGCTGGGGCCCCCAGCCACCCGTGCCAGGAGCCTT 576

Qy       2593 CTCCCCTTGCCGTTGGAGGGCAAGGCCAAACGCAGCCACTCCTTCGACATGGGGGACTTT 2652
              ||||| |||||  ||||||||||||||||||| |||||||| || |||||||||||||||
Db        575 CTCCCTTTGCCCCTGGAGGGCAAGGCCAAACGAAGCCACTCTTTTGACATGGGGGACTTT 516

Qy       2653 GCTGCTGCGGCGGCGGGAGGGGTCGTGCCGGGCGGCTACAGTCCTCCTCGGAAGGTCTCG 2712
              || ||||| || ||       |  |  || || ||||||||||| ||| ||  ||||||
Db        515 GCAGCTGCAGCTGC------AGCTGCTCCTGGTGGCTACAGTCCCCCTAGGCGGGTCTCA 462

Qy       2713 AACATCTGGACGAAGCGCAGCCTCTCTGTCAACGGCATGCTCTTGCCCTTTGAGGAGAGT 2772
              ||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||| |
Db        461 AACATCTGGACGAAGCGCAGCCTGTCTGTCAATGGCATGCTCTTGCCCTTTGAGGAGAAT 402

Qy       2773 GACCTGGTGGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATGGAGAGCACGGTC 2832
              ||||||||||| || |||||||| ||||||||||| || ||||| |||||||| || ||
Db        401 GACCTGGTGGGAGCTCGGGGGACATTTGGCAGCTCAGAGTGGGTAATGGAGAGTACTGTG 342

Qy       2833 TAGGTGGGGGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGGAGAAGGGGAAAGAATC- 2891
              |||   |||  ||     |  || | |       |      |  | || |||| || ||
Db        341 TAGCCAGGGAAGGAGTACCGTCCCTCCCGCCCTTCCTCAAAGTGGGAGAGGAAGGAGTCA 282

Qy       2892 ------TCACTGGCAAGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGACAGT 2945
                     ||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        281 GAGTCAGCACTGGCAAGTATTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGACAGT 222

Qy       2946 TTCGTCTCCCTGTCAATGGCCTCGTGTC------CCCCCCTACCCCGCAACACCCACATC 2999
               | |||||||||||||  ||||| ||||      |||  |||| || |    || |  ||
Db        221 CTTGTCTCCCTGTCAACAGCCTCATGTCCTTCTGCCCAGCTACACCCCCTTCCCTATGTC 162

Qy       3000 ACCTCCCCA-----------CCACCCGGCCGGGGTGTGCTCAGGGAATGTGGACTCGCTC 3048
              |||| ||||            |  ||   |||||||||||||||||||| ||||||||||
Db        161 ACCTTCCCAGGCCTGTCCCCACCACCCAGCGGGGTGTGCTCAGGGAATG-GGACTCGCTC 103

Qy       3049 AAATGCCGGACTGAGCCCTGAGTGTTTGGAAAGGCGAGACTCCGCCTTTCTAATCACAAA 3108
              |  || | ||||||||||||||||||||||||||| |    | ||||||||| |||||||
Db        102 ATGTGTCAGACTGAGCCCTGAGTGTTTGGAAAGGCTAACAACTGCCTTTCTATTCACAAA 43

Qy       3109 TGTAGCCTACAAGCAAGCGGCTTTGGA 3135
              ||||| | ||||||||||||||| | |
Db         42 TGTAG-CGACAAGCAAGCGGCTTCGAA 17

RESULT 14
AW161909/c
LOCUS       AW161909                 397 bp    mRNA    linear   EST 09-NOV-1999
DEFINITION  au71e01.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
            IMAGE:2781720 3', mRNA sequence.
ACCESSION   AW161909
VERSION     AW161909.1  GI:6300942
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 397)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: au71e01.y1
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 390.
FEATURES             Location/Qualifiers
     source          1..397
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2781720"
                     /clone_lib="Schneider fetal brain 00004"
                     /sex="male"
                     /tissue_type="frontal lobe"
                     /dev_stage="5 months post-conception"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pBluescript SK (Stratagene);
                     Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
                     prepared from human fetal brain tissue.  5' and 3'
                     adaptors were used in cloning as follows: 5' adaptor
                     sequence:
                     5'-GAGAGAGAGAAGAGCTCAAGGATCCTTAATTAAATTAATCCCCCCCCCCCC-3'
                     and 3' adaptor sequence:
                     5'-GAGAGAGAGACTCGAGTTTTTTTTTTTTTTTTTT-3'.  The library was
                     size-selected for >0.5 kb inserts and has an average
                     insert size estimated at 1.2 kb.  This library was
                     constructed using the CAP-trapper method for full-length
                     enrichment and has not undergone amplification.   Library
                     was constructed by Dr. Claudio Schneider (LNCIB-Area
                     Science Park,  Trieste, Italy). "
BASE COUNT       84 a    113 c     91 g    100 t      9 others
ORIGIN

  Query Match             10.0%;  Score 313;  DB 10;  Length 397;
  Best Local Similarity   90.6%;  Pred. No. 2e-56;
  Matches  328;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       2781 GGGGGCCCGGGGGACTTTTGGCAGCTCCGAATGGGTGATGGAGAGCACGGTCTAGGTGGG 2840
              |||| || ||||| |||||||||||||| ||||||||||||| ||| |||||||||||||
Db        395 GGGGCCCGGGGGGTCTTTTGGCAGCTCCCAATGGGTGATGGAAAGCCCGGTCTAGGTGGG 336

Qy       2841 GGTGGGCATGCTCCCTTTCCTGTGCGCAGGGTGGGAGAAGGGGAAAGAATCTCACTGGCA 2900
              |||||||| ||||||||||||||||||||||||||| ||||||||| | | |||||||||
Db        335 GGTGGGCAAGCTCCCTTTCCTGTGCGCAGGGTGGGAAAAGGGGAAANATTTTCACTGGCA 276

Qy       2901 AGTGTTTGTGGAGTTTCCATGGTGATGTTTACATCCAGGGACAGTTTCGTCTCCCTGTCA 2960
              || |||||||||||||||||||||||||||||| |||||||| |||| ||||||| |  |
Db        275 AGGGTTTGTGGAGTTTCCATGGTGATGTTTACAACCAGGGACGGTTTTGTCTCCCNGNNA 216

Qy       2961 ATGGCCTCGTGTCCCCCCCTACCCCGCAACACCCACATCACCTCCCCACCACCCGGCCGG 3020
              | ||||| ||||||||||  |||| | |||||||||||| ||||||| || |||||||||
Db        215 AGGGCCTNGTGTCCCCCCNNACCCNGNAACACCCACATCCCCTCCCCCCCCCCCGGCCGG 156

Qy       3021 GGTGTGCTCAGGGAATGTGGACTCGCTCAAATGCCGGACTGAGCCCTGAGTGTTTGGAAA 3080
              ||||||||||||||||||||| |||||||||||||||| |||||||||||||||||||||
Db        155 GGTGTGCTCAGGGAATGTGGATTCGCTCAAATGCCGGATTGAGCCCTGAGTGTTTGGAAA 96

Qy       3081 GGCGAGACTCCGCCTTTCTAATCACAAATGTAGCCTACAAGCAAGCGGCTTTGGATTGCT 3140
              || |||||||| ||||| ||||||||||||||||||||||||||| |||||||||||| |
Db         95 GGGGAGACTCCCCCTTTTTAATCACAAATGTAGCCTACAAGCAAGGGGCTTTGGATTGTT 36

Qy       3141 TA 3142
              ||
Db         35 TA 34

RESULT 15
BM046297
LOCUS       BM046297                 708 bp    mRNA    linear   EST 07-NOV-2001
DEFINITION  603626095F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5452630 5',
            mRNA sequence.
ACCESSION   BM046297
VERSION     BM046297.1  GI:16775564
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 708)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1944  row: n  column: 23
            High quality sequence stop: 704.
FEATURES             Location/Qualifiers
     source          1..708
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5452630"
                     /clone_lib="NIH_MGC_40"
                     /tissue_type="carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed  by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California,  Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life  Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      101 a    272 c    209 g    126 t
ORIGIN

  Query Match              9.9%;  Score 310.6;  DB 13;  Length 708;
  Best Local Similarity   65.8%;  Pred. No. 7e-56;
  Matches  473;  Conservative    0;  Mismatches  234;  Indels   12;  Gaps    1;

Qy          651 CTTCATCATCCACATCAGCCGCCAGGACTTTGCCAACATGACGGGGCTGGTGGACCTGAC 710
                |||||||     | |  | | ||  |||||   |||||||||||| ||||||||||||||
Db            2 CTTCATCCAGGCCCTGGGGCCCCCTGACTTCCGCAACATGACGGGACTGGTGGACCTGAC 61

Qy          711 CCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGAGCCT 770
                 |||||  | ||  |||||| || |||   | ||    ||||||  |||||||||||||||
Db           62 ACTGTCTCGCAATGCCATCACCCGCATTGGGGCCCGCGCCTTTGGGGACCTCGAGAGCCT 121

Qy          771 CCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCTCCGGGG 830
                 || ||||| || |||||| ||||  |||||      || ||    |  | |||||||||
Db          122 GCGTTCCCTCCACCTTGACGGCAACAGGCTGGTGGAGCTGGGCACCGGGAGCCTCCGGGG 181

Qy          831 CCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACCAGCTGGGCGGCATCGCAGATGA 890
                ||  ||||| ||||||||||| ||| | | |  ||||||||||||| |||||||    | 
Db          182 CCCCGTCAATCTGCAGCACCTCATCCTCAGCGGCAACCAGCTGGGCCGCATCGCGCCGGG 241

Qy          891 GGCTTTTGAGGACTTCCTGCTGACATTGGAGGATCTGGACCTCTCCTACAACAACCTCCA 950
                 || || || ||||||||   ||   |||||||| |||||||| ||||||||||||||||| 
Db          242 AGCCTTCGACGACTTCCTAGAGAGCCTGGAGGACCTGGACCTGTCCTACAACAACCTCCG 301

Qy          951 TGGCCTGCCGTGGGACTCCGTGCGACGCATGGTCAACCTCCACCAGCTGAGCCTGGACCA 1010
                     |||| |||| |  | |  |   |||||     |||||||   || | |||||||||
Db          302 GCAGGTGCCCTGGGCCGGCATCGGCGCCATGCCTGCCCTGCACACCCTCAACCTGGACCA 361

Qy         1011 CAACCTGCTGGATCACATCGCCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCCT 1070
                 |||||  | ||     ||   |    ||| ||  | ||     | ||  |||||||
Db          362 TAACCTTATTGACGCACTGCCCCCAGGCGCCTTCGCCCAGCTCGGTCAGCTCTCCCGCCT 421

Qy         1071 GGATCTCACCTCCAATCGGCTGCAGAAGCTGCCCCCTGATCCCATCTTTGCCCGCTCCCA 1130
                ||| |||||||||||| || |||   | ||||| | || || ||  | ||  | ||    | 
Db          422 GGACCTCACCTCCAACCGCCTGGCCACGCTGGCTCCGGACCCGCTTTTCTCTCGTGGGCG 481

Qy         1131 GGCTTCGGCTTTGACAGCCACACCCTTTGCCCCACCCTTGTCCTTTAGTTTTGGGGGTAA 1190
                 | | | |               ||| | || | ||| ||    | || ||| | || ||
Db          482 TGATGCAGAGG------------CCTCTCCCGCCCCCCTGGTGCTGAGCTTTAGCGGGAA 529

Qy         1191 CCCACTTCACTGCAATTGTGAGCTTCTCTGGCTGCGGAGGCTCGAGCGGGACGATGACCT 1250
                ||| || |||||||| |||||||| || ||||||||| |||| | ||||    || |||||
Db          530 CCCCCTGCACTGCAACTGTGAGCTGCTGTGGCTGCGGCGGCTGGCGCGGCCGGACGACCT 589

Qy         1251 GGAAACCTGTGGCTCCCCAGGGGGCCTCAAGGGTCGCTACTTCTGGCATGTGCGTGAGGA 1310
                |||||| || | ||||||    |||||     || ||||||||||||   ||   |||| 
Db          590 GGAAACGTGCGCCTCCCCGACCGGCCTGGCCGGCCGCTACTTCTGGGCAGTGGCCGAGGG 649

Qy         1311 GGAGTTTGTGTGCGAGCCGCCTCTCATCACCCAGCACACACACAAGTTGCTGGTTCTGG 1369
                 |||||    || |||||  | |||||  |||  |||||  ||     |   ||| || |
Db          650 CGAGTTCTCCTGTGAGCCTGCCCTCATTGCCCGACACACGCAGCGCCTCTGGGTGCTTG 708

Search completed: June  8, 2003, 19:02:56
Job time : 3928 secs
```

```
                  GenCore version 5.1.6
         Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         June  8, 2003, 19:03:03 ; Search time 63 Seconds
                                          (without alignments)
                                          1668.805 Million cell updates/sec

Title:          US-09-831-846-2
Perfect score:  4094
Sequence:       1 METLLGGLLAFGMAFAVVDA..........DLVGARGTFGSSEWVMESTV 789

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:      908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_101002:*
                1:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                         %
Result                 Query
  No.     Score    Match Length DB  ID                    Description
 -----------------------------------------------------------------------
    1      4094    100.0    789  21  AAB09968              Human brain-specif
    2      4087     99.8    789  22  AAU28092              Novel human secret
    3      4087     99.8    789  22  AAM39059              Human polypeptide
    4      4075     99.5    785  21  AAB12448              Human hh00149 prot
    5      1846     45.1    766  22  AAG67505              Amino acid sequenc
    6    1820.5     44.5    719  23  ABP43533              Human secreted pro
    7    1703.5     41.6    771  22  AAG67512              Amino acid sequenc
    8    1667.5     40.7    635  23  AAE23980              Human LP220 secret
    9      1656     40.4    322  22  ABG13536              Novel human diagno
   10    1586.5     38.8    636  22  AAU32870              Novel human secret
   11      1573     38.4    628  22  AAB84469              Amino acid sequenc
   12      1573     38.4    628  22  AAG67523              Amino acid sequenc
   13      1566     38.3    628  22  AAG65805              Human leucine-rich
   14    1563.5     38.2    627  23  ABG34079              Human Pro peptide
   15    1545.5     37.8    551  23  AAE17484              Human leucine-rich
   16    1468.5     35.9    526  22  ABG04827              Novel human diagno
   17    1319.5     32.2    533  22  AAG67511              Amino acid sequenc
   18    1313.5     32.1    503  22  ABG05946              Novel human diagno
   19      1241     30.3    768  22  ABG19413              Novel human diagno
   20      1239     30.3    479  22  AAG67522              Amino acid sequenc
   21      1144     27.9    468  22  AAB70072              Human secreted pro
   22      1144     27.9    468  23  ABG65511              Human albumin fusi
   23       721     17.6    160  22  AAM40845              Human polypeptide
   24       651     15.9    136  22  AAU28280              Novel human secret
   25       511     12.5    180  22  ABB03072              Human expressed po
   26       511     12.5    180  22  AAU20510              Human secreted pro
   27       452     11.0    713  23  AAU91335              Human novel secret
   28       451     11.0    713  22  AAE13006              Human leucine-rich
   29     443.5     10.8    656  23  AAU91341              Human novel secret
   30     435.5     10.6    640  20  AAY08100              Human PRO331 prote
   31     435.5     10.6    640  20  AAW85722              Novel protein (Clo
   32     435.5     10.6    640  20  AAY13394              Amino acid sequenc
   33     435.5     10.6    640  21  AAB24407              Human PRO331 prote
   34     435.5     10.6    640  21  AAY70673              Human PRO331 prote
   35     435.5     10.6    640  22  AAU12355              Human PRO331 polyp
   36     435.5     10.6    640  22  AAU00826              Human immune respo
   37     435.5     10.6    640  22  AAB80262              Human PRO331 prote
   38     435.5     10.6    640  22  AAB65292              Human PRO331 prote
   39     435.5     10.6    640  22  AAB53089              Human angiogenesis
   40     420.5     10.3    606  23  ABG60230              Human Slit-3-like
   41     420.5     10.3    606  23  AAU85411              Human protein NOV9
   42       419     10.2    606  23  ABG60229              Human Slit-3-like
   43       419     10.2    606  23  AAU85410              Human protein NOV8
   44     418.5     10.2    579  22  AAE06804              Mature human neuro
   45     418.5     10.2    606  22  AAE06799              Human neuronal gui

                                  ALIGNMENTS

RESULT 1
AAB09968
ID   AAB09968 standard; Protein; 789 AA.
XX
AC   AAB09968;
XX
DT   19-OCT-2000  (first entry)
XX
DE   Human brain-specific transmembrane glycoprotein.
XX
KW   Transmembrane glycoprotein; human; brain; cellular signal transducer;
KW   neuroactive; neuroprotective; cerebroprotective; drug development;
KW   treatment; nervous disease; diagnostic.
XX
OS   Homo sapiens.
XX
PN   WO200031256-A1.
XX
PD   02-JUN-2000.
XX
PF   18-NOV-1999;   99WO-JP06449.
XX
PR   20-NOV-1998;   98JP-0331727.
XX
PA   (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
PA   (KAZU-) KAZUSA DNA RES INST.
XX
PI   Funahashi S,  Miyata S,  Nomura N,  Nagase T,  Ohara O;
XX
DR   WPI; 2000-411954/35.
DR   N-PSDB; AAA40083.
XX
PT   Gene encoding a brain-specific transmembrane glycoprotein that has a
```

```
PT   typical PDZ protein binding motif and functions as a cellular signal
PT   transducer, useful in developing drugs for treating nervous diseases -
XX
PS   Claim 1a; Page 52-57; 63pp; Japanese.
XX
CC   This invention describes a novel gene encoding a human brain-specific
CC   transmembrane glycoprotein that has a typical PDZ protein binding motif
CC   and functions as a cellular signal transducer. The product of the
CC   invention has neuroactive, neuroprotective and cerebroprotective
CC   activity. The gene and encoded protein are useful in developing drugs
CC   for treating nervous diseases, and also for studying functions of the
CC   nervous system or onset mechanism of nerve-related diseases. They may
CC   also be used for the production of diagnostic reagents such as
CC   oligonucleotide probes and antibodies for detecting proteins comprising
CC   PDZ domains and the nucleic acids that encode them. They may also be
CC   used to produce and identify modulators of the proteins expression and
CC   activity such as antisense sequences and antibodies. This sequence
CC   represents the human brain-specific transmembrane glycoprotein described
CC   in the method of the invention.
XX
SQ   Sequence   789 AA;

  Query Match            100.0%;  Score 4094;  DB 21;  Length 789;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-303;
  Matches  789;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360

Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600

Qy        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660

Qy        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720

Qy        721 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780

Qy        781 SEWVMESTV 789
              |||||||||
Db        781 SEWVMESTV 789


RESULT 2
AAU28092
ID   AAU28092 standard; Protein; 789 AA.
XX
AC   AAU28092;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secretory protein, Seq ID No 261.
XX
KW   Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW   ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW   transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW   amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW   ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW   gut protection; lung; liver fibrosis; immune deficiency; infection;
KW   severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW   multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW   fertility; analgesic; pain; antigen.
XX
OS   Homo sapiens.
XX
PN   WO200166689-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US04942.
XX
PR   07-MAR-2000; 2000US-0519705.
PR   19-MAY-2000; 2000US-0574454.
PR   17-JUN-2000; 2000US-0596193.
PR   14-JUL-2000; 2000US-0616847.
PR   19-SEP-2000; 2000US-0665363.
PR   20-OCT-2000; 2000US-0693267.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Xu C,  Wehrman T,  Ren F,  Ma Y,  Zhou P;
PI   Zhao QA,  Yang Y,  Drmanac RT,  Zhang J,  Chen R,  Xue AJ,  Wang J;
XX
DR   WPI; 2001-589934/66.
DR   N-PSDB; AAS44992.
XX
PT   Novel polypeptides and nucleic acids obtained from cDNA libraries
PT   prepared from various human tissues, for diagnosis and treatment of
PT   cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS   Example 4; SEQ ID No 261; 107pp; English.
XX
CC   The invention relates to novel isolated human secreted polypeptides (I)
CC   and polynucleotides (II). (I) and (II) are useful for treating
CC   inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC   ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC   involved in increasing haematopoiesis, stem cell survival, bone growth
CC   and remodeling. (I), (II) and modulators of (II) are useful for
CC   prophylaxis or treatment of one or more cancers. (II) is also useful for
CC   creating transgenic animals useful for studying the in vivo activities of
CC   the polypeptide as well as for studying modulators of the polypeptides.
CC   (I) induces the proliferation of neural cells and regeneration of nerve
CC   and brain tissue and is useful for the treatment of central and
CC   peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC   Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC   sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
```

```
CC   activity, regulation of haematopoiesis and is useful for treating myeloid
CC   or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC   and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC   tissue growth, and in tissue repair, healing of burns, incisions,
CC   ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC   disorders, or periodontal disease. Furthermore, (I) is also useful for
CC   gut protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues, various immune deficiencies and
CC   disorders including severe combined immunodeficiency (SCID), bacterial or
CC   fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC   rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC   reactions and conditions, such as asthma or other respiratory problems.
CC   In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC   fertility, metabolism, catabolism, anabolism, storage or elimination of
CC   dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC   analgesic effects or other pain reducing effects, immunoglobulin like
CC   activity and can act as an antigen in a vaccine composition to raise an
CC   immune response. AAU28020-AAU28395 represent novel human secreted protein
CC   amino acid sequences of the invention.
XX
SQ   Sequence   789 AA;

  Query Match             99.8%;  Score 4087;  DB 22;  Length 789;
  Best Local Similarity   99.9%;  Pred. No. 6e-303;
  Matches  788;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    .0;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTPDIFITTSQDSGAFTCI 360

Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600

Qy        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660

Qy        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720

Qy        721 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780

Qy        781 SEWVMESTV 789
              |||||||||
Db        781 SEWVMESTV 789


RESULT 3
AAM39059
ID   AAM39059 standard; Protein; 789 AA.
XX
AC   AAM39059;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 2204.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI58215.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2204; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
```

```
CC   specification.
XX
SQ   Sequence   789 AA;

  Query Match             99.8%;  Score 4087;  DB 22;  Length 789;
  Best Local Similarity   99.9%;  Pred. No. 6e-303;
  Matches  788;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTPDIFITTSQDSGAFTCI 360

Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600

Qy        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660

Qy        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720

Qy        721 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780

Qy        781 SEWVMESTV 789
              |||||||||
Db        781 SEWVMESTV 789


RESULT 4
AAB12448
ID   AAB12448 standard; Protein; 785 AA.
XX
AC   AAB12448;
XX
DT   19-OCT-2000  (first entry)
XX
DE   Human hh00149 protein SEQ ID NO:4.
XX
KW   Human; ubiquitin-like protein; 149Y2H#151; hh00149; brain;
KW   2-hybrid screening; neuroprotective; signal transducer;
KW   nervous system disease; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO200031255-A1.
XX
PD   02-JUN-2000.
XX
PF   18-NOV-1999;   99WO-JP06448.
XX
PR   20-NOV-1998;   98JP-0331701.
XX
PA   (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI   Funahashi S,  Miyata S;
XX
DR   WPI; 2000-400066/34.
DR   N-PSDB; AAA60605.
XX
PT   Gene encoding ubiquitin-like protein which interacts with protein
PT   hh00149, useful in the diagnosis and treatment of diseases associated
PT   with the nervous system -
XX
PS   Example 2; Page 74-79; 88pp; Japanese.
XX
CC   The present invention describes a ubiquitin-like protein, designated
CC   149Y2H#151, which interacts with protein hh00149 expressed specifically
CC   in the brain. The 149Y2H#151 protein has neuroprotective activity, and
CC   is a signal transducer. The 149Y2H#151 gene and encoded protein are
CC   useful in the diagnosis and treatment of diseases associated with the
CC   nervous system. The protein can interact with protein hh00149 expressed
CC   specifically in the brain. The present sequence represents the human
CC   hh00149 protein.
XX
SQ   Sequence   785 AA;

  Query Match             99.5%;  Score 4075;  DB 21;  Length 785;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-302;
  Matches  785;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFI 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFI 60

Qy         65 IHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLV 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLV 120

Qy        125 NLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNL 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNL 180

Qy        185 LDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPL 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPL 240

Qy        245 HCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQ 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQ 300

Qy        305 AATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANA 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANA 360

Qy        365 AGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPER 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPER 420

Qy        425 AVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSG 484
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSG 480

Qy        485 TGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVA 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVA 540

Qy        545 TLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPPKVVV 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPPKVVV 600

Qy        605 RNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAFASLD 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 RNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAFASLD 660

Qy        665 LKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKRSHSF 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKRSHSF 720

Qy        725 DMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSEWV 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 DMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSEWV 780

Qy        785 MESTV 789
              |||||
Db        781 MESTV 785


RESULT 5
AAG67505
ID   AAG67505 standard; Protein; 766 AA.
XX
AC   AAG67505;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human secreted polypeptide.
XX
KW   Human; secreted polypeptide; nervous disease; muscular disease; tumour;
KW   gastrointestinal ulceration; spinal cord disease; trachea disease;
KW   thyroid gland disease; ovary disease; prostate disease; heart disease;
KW   renal gland disease; small intestine disease; thymus disease;
KW   lymph node disease; muscular system disease; colon disease;
KW   lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
KW   myocardial infarction; angioplasty; liver disease; coagulation disorder;
KW   microbial disease; immune disorder; inflammation; transplant rejection;
KW   bone thickness; bone density; ferroxidase loss; apoptosis;
KW   vascular smooth cell proliferation; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200166690-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US07143.
XX
PR   06-MAR-2000; 2000US-0187107.
PR   13-MAR-2000; 2000US-0188916.
PR   03-OCT-2000; 2000US-0236874.
PR   03-OCT-2000; 2000US-0237846.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Agarwal P,  Murdoch PR,  Rizvi SK,  Smith RF,  Xiang Z;
XX
DR   WPI; 2001-570768/64.
DR   N-PSDB; AAH78197.
XX
PT   Novel isolated secreted polypeptide useful for treating nervous and
PT   muscular diseases, gastrointestinal ulceration, coagulation and immune
PT   disorders, microbial diseases, inflammation and transplant rejection -
XX
PS   Claim 1; Page 58-60; 102pp; English.
XX
CC   The present sequence represents a human secreted polypeptide. The
CC   secreted polypeptides and polynucleotides are useful for treating
CC   nervous and muscular diseases, for inhibiting tumour formation and
CC   metastasis, for treating gastrointestinal ulceration, for preventing
CC   and treating diseases in spinal cord, thyroid gland, ovary, prostate,
CC   renal gland, small intestine, heart, trachea, thymus, lymph node,
CC   muscular system and colon, for treating lipase deficiency in cystic
CC   fibrosis and pancreatitis, for treating undesirable clot formation
CC   such as myocardial infarction, during angioplasty and all surgical
CC   procedures that require decreased blood clot formation, for treating
CC   liver diseases, coagulation disorders and microbial diseases, for
CC   treating immune disorders, for treating inflammation and transplant
CC   rejection, for enhancing bone thickness and increasing bone density,
CC   for reducing the loss of essential ferroxidases, for suppressing
CC   apoptosis, and for regulating vascular smooth cell proliferation. They
CC   may also be used as vaccines.
XX
SQ   Sequence   766 AA;

  Query Match             45.1%;  Score 1846;  DB 22;  Length 766;
  Best Local Similarity   48.5%;  Pred. No. 5.3e-132;
  Matches  383;  Conservative  128;  Mismatches  225;  Indels   54;  Gaps   11;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              || :|  |   |:|      ||| |||| || :| |||  ||||||||:||||||||||
Db          1 MEKVLFYLFLIGIA-VKAQICPKRCVCQILSPNLATLCAKKGLLFVPPNIDRRTVELRLA 59

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
               ||: :| |:|||||| |||||||||||| | | :| || :||:|||:||||  :  |
Db         60 DNFVTNIKRKDFANMTSLVDLTLSRNTISFITPHAFADLRNLRALHLNSNRLTKITNDMF 119

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
               || || |||:|||||  |:  ||:| :  ||:||||||||  :|||:| :||:|| |||
Db        120 SGLSNLHHLILNNNQLTLISSTAFDD-VFALEELDLSYNNLETIPWDAVEKMVSLHTLSL 178

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              |||::|:| :|||: | |: |||:|||:||||||||:| |:|  | :        : |||
Db        179 DHNMIDNIPKGTFSHLHKMTRLDVTSNKLQKLPPDPLFQRAQVLATSGIISPSTFALSFG 238

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              |||||||||||||||| |:|||||| ||  | ||||| : ||||:|||||||:|||:: |
Db        239 GNPLHCNCELLWLRRLSREDDLETCASPPLLTGRYFWSIPEEEFLCEPPLITRHTHEMRV 298

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              |||| |||:||| ||| | |||::|: :|: |::|: ||||||||| ||| :|:||||||
Db        299 LEGQRATLRCKARGDPEPAIHWISPEGKLISNATRSLVYDNGTLDILITTVKDTGAFTCI 358

Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              |:| ||||| :|:: |::|||| |||:    |    |||: |:|:      | |
Db        359 ASNPAGEATQIVDLHIIKLPHLLNSTNHIHEPDPGSSDISTSTKSGSNTSSSNG---DTK 415

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
                :  ::|:| |:::||:|::  :: | ::|:|:||| : |: |:||||| ::| |:|||
Db        416 LSQDKIVVAEATSSTALLKFNFQRNIPGIRMFQIQYNGTYDDTLVYRMIPPTSKTFLVNN 475

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              | :|| |||||||::||  |:|||| :||| || |: || :|  | || ||||||::|||
Db        476 LAAGTMYDLCVLAIYDDGITSLTATRVVGCIQFTTEQDYVRCHFMQSQFLGGTMIIIIGG 535

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600
              ||||::||||:|||:||||||:    |:   ||||||||||||        :   ||
Db        536 IIVASVLVFIIILMIRYKVCNNNGQHKV-TKVSNVYSQTNGAQ----IQGCSVTLPQSVS 590

Qy        601 KVVVRNELLDFTASLARASDSS---SSSSLGSGEAAGLGRAPWRIPPS------------ 645
              |  | :|     |   :|:  :   || :  | :::    |   :|||
Db        591 KQAVGHE---ENAQCCKATSDNVIQSSETCSSQDSSTTTSA---LPPSWTSSTSVSQKQK 644

Qy        646 -APRPKPSLDRLMGAFASLDLKSQRKEE----LLDSRTPAGRGAGTSARGHHSDREPLLG 700
                   ||| :    |  ::: ::  :       | || |     | :::  |      :
Db        645 RKTGTKPSTEPQNEAVTNVESQNTNRNNSTALQLASRPPDSVTEGPTSKRAHIKPSKFIT 704
```

```
Qy        701 PPAARARSLLPLPLEGKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVN 760
               || |: :     | |: |  :                   |        :: |||:|:|
Db        705 LPAERSGARHKYSLNGELKEYYC------------------YINSPNTCGLFPKRSMSMN 746

Qy        761 GMLLPFEESD 770
               | :  : ||
Db        747 VMFIQSDCSD 756

RESULT 6
ABP43533
ID   ABP43533 standard; Protein; 719 AA.
XX
AC   ABP43533;
XX
DT   08-AUG-2002  (first entry)
XX
DE   Human secreted protein (SCEP) 57.
XX
KW   Human; secreted protein; SECP; SECP expression; gene therapy;
KW   protein therapy: immune system disorders; AIDS; thymic hypoplasia;
KW   anaemia; asthma; Crohn's disease; neurological disorder; epilepsy;
KW   Huntington's disease; dementia; Parkinson's disease; Down's syndrome;
KW   developmental disorder; cell proliferative disorder; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200226982-A2.
XX
PD   04-APR-2002.
XX
PF   25-SEP-2001; 2001WO-US30042.
XX
PR   29-SEP-2000; 2000US-236869P.
PR   11-OCT-2000; 2000US-239812P.
PR   12-OCT-2000; 2000US-240108P.
PR   17-OCT-2000; 2000US-241282P.
PR   20-OCT-2000; 2000US-242218P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Tang YT,  Nguyen DB,  Yao MG,  Xu Y,  Tribouley CM;
PI   Sanjanwala MS,  Walia NK,  Baughn MR,  Sapperstein SK,  Lal P;
PI   Thornton M,  Gandhi AR,  Ramkumar J,  Elliott VS,  Arvizu C;
PI   Thangavelu K,  Gietzen KJ,  Ding L,  Au-young J,  Tran B,  Policky JL;
PI   Lee S,  Lu DAM,  Burford N,  Warren BA,  Gururajan R,  Duggan BM;
PI   Honchell CD,  Hafalia AJA;
XX
DR   WPI; 2002-394239/42.
DR   N-PSDB; ABN99416.
XX
PT   New human secreted proteins, useful for diagnosing, treating or
PT   preventing immune system disorders (e.g. Crohn's disease), neurological
PT   disorders (e.g. Parkinson's disease), or cell proliferative disorders
PT   (e.g. cancers)  -
XX
PS   Claim 1; Page 188-189; 238pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of human
CC   secreted proteins (SECP). The SECP DNA and amino acid sequences of the
CC   invention are useful for treating/preventing disorders associated with
CC   decreased or elevated expression of SECP. The SECP DNA and protein
CC   sequences are specifically useful for treating/preventing (i.e. gene
CC   therapy and protein therapy): immune system disorders (e.g. AIDS, thymic
CC   hypoplasia, anaemia, asthma or Crohn's disease); neurological disorders
CC   (e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease);
CC   developmental disorders (e.g. Down's syndrome); and cell proliferative
CC   disorders (e.g. cancer). The proteins ABP43477 - ABP43543 represent the
CC   human secreted proteins (SECP) of the invention.
XX
SQ   Sequence   719 AA;
```

```
  Query Match             44.5%;  Score 1820.5;  DB 23;  Length 719;
  Best Local Similarity   52.1%;  Pred. No. 4.3e-130;
  Matches  371;  Conservative  114;  Mismatches  190;  Indels   37;  Gaps   11;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              || :|  |   |:|      ||| |||| || :| |||  ||||||||:||||||||||
Db          1 MEKILFYLFLIGIA-VKAQICPKRCVCQILSPNLATLCAKKGLLFVPPNIDRRTVELRLA 59

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
               ||: :| |:|||||| |||||||||||| | | :| || :||:|||:||||  :  |
Db         60 DNFVTNIKRKDFANMTSLVDLTLSRNTISFITPHAFADLRNLRALHLNSNRLTKITNDMF 119

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
               || || |||:|||||  |:  ||:| :  ||:||||||||  :|||:| :||:|| |||
Db        120 SGLSNLHHLILNNNQLTLISSTAFDD-VFALEELDLSYNNLETIPWDAVEKMVSLHTLSL 178

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              |||::|:| :|||: | |: |||:|||:||||||||:| |:|  | :         :|||
Db        179 DHNMIDNIPKGTFSHLHKMTRLDVTSNKLQKLPPDPLFQRAQVLATSGIISPSTFALSFG 238

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              |||||||||||||||| |:|||||| ||  | ||||| : ||||:|||||||:|||:: |
Db        239 GNPLHCNCELLWLRRLSREDDLETCASPPLLTGRYFWSIPEEEFLCEPPLITRHTHEMRV 298

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              |||| |||:||| ||| | |||::|: :|: |::|: ||||||||| ||| :|:||||||
Db        299 LEGQRATLRCKARGDPEPAIHWISPEGKLISNATRSLVYDNGTLDILITTVKDTGAFTCI 358

Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              |:| ||||| :|:: |::|||| |||:    |    |||: |:|:      | |
Db        359 ASNPAGEATQIVDLHIIKLPHLLNSTNHIHEPDPGSSDISTSTKSGSNTSSSNG---DTK 415

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
                :  ::|:| |:::||:|::  :: | ::|:|:||| : |: |:||||| ::| |:|||
Db        416 LSQDKIVVAEATSSTALLKFNFQRNIPGIRMFQIQYNGTYDDTLVYRMIPPTSKTFLVNN 475

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              | :|| |||||||::||  |:|||| :||| || |: || :|  | || ||||||::|||
Db        476 LAAGTMYDLCVLAIYDDGITSLTATRVVGCIQFTTEQDYVRCHFMQSQFLGGTMIIIIGG 535

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600
              ||||::||||:|||:||||||:    |:   |||||||||||||        :   ||
Db        536 IIVASVLVFIIILMIRYKVCNNNGQHKV-TKVSNVYSQTNGAQ----IQGCSVTLPQSVS 590

Qy        601 KVVVRNELLDFTASLARASDSS---SSSSLGSGEAAGLGRAPWRIPPS------------ 645
              |  | :|     |   :|:  :   || :  | :::    |   :|||
Db        591 KQAVGHE---ENAQCCKATSDNVIQSSETCSSQDSSTTTSA---LPPSWTSSTSVSQKQK 644

Qy        646 -APRPKPSLDRLMGAFASLDLKSQRKEE----LLDSRTPAGRGAG-TSARGH 691
                   ||| :    |  ::: ::  :       | || |     | || | |
Db        645 RKTGTKPSTEPQNEAVTNVESQNTNRNNSTALQLASRPPDSVTEGPTSKRAH 696


RESULT 7
AAG67512
ID   AAG67512 standard; Protein; 771 AA.
XX
AC   AAG67512;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human secreted polypeptide.
XX
KW   Human; secreted polypeptide; nervous disease; muscular disease; tumour;
KW   gastrointestinal ulceration; spinal cord disease; trachea disease;
KW   thyroid gland disease; ovary disease; prostate disease; heart disease;
KW   renal gland disease; small intestine disease; thymus disease;
KW   lymph node disease; muscular system disease; colon disease;
KW   lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
KW   myocardial infarction; angioplasty; liver disease; coagulation disorder;
KW   microbial disease; immune disorder; inflammation; transplant rejection;
KW   bone thickness; bone density; ferroxidase loss; apoptosis;
KW   vascular smooth cell proliferation; vaccine.
```

```
XX
OS   Homo sapiens.
XX
PN   WO200166690-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US07143.
XX
PR   06-MAR-2000; 2000US-0187107.
PR   13-MAR-2000; 2000US-0188916.
PR   03-OCT-2000; 2000US-0236874.
PR   03-OCT-2000; 2000US-0237846.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Agarwal P,  Murdoch PR,  Rizvi SK,  Smith RF,  Xiang Z;
XX
DR   WPI; 2001-570768/64.
DR   N-PSDB; AAH78204.
XX
PT   Novel isolated secreted polypeptide useful for treating nervous and
PT   muscular diseases, gastrointestinal ulceration, coagulation and immune
PT   disorders, microbial diseases, inflammation and transplant rejection -
XX
PS   Claim 1; Page 69-72; 102pp; English.
XX
CC   The present sequence represents a human secreted polypeptide. The
CC   secreted polypeptides and polynucleotides are useful for treating
CC   nervous and muscular diseases, for inhibiting tumour formation and
CC   metastasis, for treating gastrointestinal ulceration, for preventing
CC   and treating diseases in spinal cord, thyroid gland, ovary, prostate,
CC   renal gland, small intestine, heart, trachea, thymus, lymph node,
CC   muscular system and colon, for treating lipase deficiency in cystic
CC   fibrosis and pancreatitis, for treating undesirable clot formation
CC   such as myocardial infarction, during angioplasty and all surgical
CC   procedures that require decreased blood clot formation, for treating
CC   liver diseases, coagulation disorders and microbial diseases, for
CC   treating immune disorders, for treating inflammation and transplant
CC   rejection, for enhancing bone thickness and increasing bone density,
CC   for reducing the loss of essential ferroxidases, for suppressing
CC   apoptosis, and for regulating vascular smooth cell proliferation. They
CC   may also be used as vaccines.
XX
SQ   Sequence  771 AA;

  Query Match             41.6%;  Score 1703.5;  DB 22;  Length 771;
  Best Local Similarity   46.5%;  Pred. No. 4e-121;
  Matches  376;  Conservative  110;  Mismatches  213;  Indels  109;  Gaps   18;

Qy          21 CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
               ||  |:|||:: :|  ||   ||||||| |||| |||||  |||  : |:|||||| ||
Db          34 CPGRCICQNVAPTLTMLCAKTGLLFVPPAIDRRVVELRLTDNFIAAVRRRDFANMTSLVH 93

Qy          81 LTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIA 140
               ||||||||  :   :| || :||:||||||||  :  | |||| ||:|||: |||:  :
Db          94 LTLSRNTIGQVAAGAFADLRALRALHLDSNRLAEVRGDQLRGLGNLRHLILGNNQIRRVE 153

Qy         141 DEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLA 200
                 ||: || |:||||||||||  |||::| :||||: |:|||||:||||||||  | ||
Db         154 SAAFDAFLSTVEDLDLSYNNLEALPWEAVGQMVNLNTLTLDHNLIDHIAEGTFVQLHKLV 213

Qy         201 RLDLTSNRLQKLPPDPIFARSQASALTATPFAP-PLSFSFGGNPLHCNCELLWLRRLERD 259
               |||:||||| ||||| :| ||| :     |  | ||: ||||||||||||||||||| |:
Db         214 RLDMTSNRLHKLPPDGLFLRSQGTG----PKPPTPLTVSFGGNPLHCNCELLWLRRLTRE 269

Qy         260 DDLETCGSPGGLKGRYFWHVREEEFVCEPPLIT-QHTHKLLVLEGQAATLKCKAIGDPSP 318
               |||||| :|  |  |||| : ||||:||||||| |   : ||:|||| :|:|:|:||| |
Db         270 DDLETCATPEHLTDRYFWSIPEEEFLCEPPLITRQAGGRALVVEGQAVSLRCRAVGDPEP 329

Qy         319 LIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMVEVSIVQ 378
               ::|||||| ||:|||||| |  :||||: ||| :||| |||||:|||||||| ||| :|
Db         330 VVHWVAPDGRLLGNSSRTRVRGDGTLDVTITTLRDSGTFTCIASNAAGEATAPVEVCVVP 389

Qy         379 LPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALV 438
               || :       |||  :    :|    :     |  |     |  || :: :|:|: | |:
Db         390 LPLM-------APPPAAPPPLTEPGSSDIATPGRPGAN--DSAAERRLVAAELTSNSVLI 440

Qy         439 KWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDT 498
               :|   :  | ::|||:||| | |: |:|||||:::: |:||:| :|  |||||||::||
Db         441 RWPAQRPVPGIRMYQVQYNSSVDDSLVYRMIPSTSQTFLVNDLAAGRAYDLCVLAVYDDG 500

Qy         499 ATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRYK 558
               || | || :||| || |  |   |: : :  ||||||: |||:|||::|||||:||:|||
Db         501 ATALPATRVVGCVQFTTAGDPAPCRPLRAHFLGGTMIIAIGGVIVASVLVFIVLLMIRYK 560

Qy         559 V-----CNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP-------------QGPP 600
               |           |:    ||:| ||||||      :| | | |             |  |
Db         561 VYGDGDSRRVKGSRSLPRVSHVCSQTNGA---GTGAAQAPALPAQDHYEALREVESQAAP 617

Qy         601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
                | |           |:| :: ::|:                      |:  | | :|
Db         618 AVAVE----------AKAMEAETASA---------------------EPEVVLGRSLGGS 646

Qy         661 A-SLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLP----LE 715
               | || |    :    :||         :| |    |   | || :   :|  :|
Db         647 ATSLCLLPSEETSGEESR---------AAVGPRRSRSGALEPPTSAPPTLALVPGGAAAR 697

Qy         716 GKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGAR 775
                : :: :||| ||:      | :   :| ||:     |||  |      |  :    ||
Db         698 PRPQQRYSFD-GDY------GALFQSHSYPRRARR--TKRHRST-----PHLDGAGGGAA 743

Qy         776 G--------------TFGSSEWVMESTV 789
               |               | |:||::||||
Db         744 GEDGDLGLGSARACLAFTSTEWMLESTV 771


RESULT 8
AAE23980
ID   AAE23980 standard; Protein; 635 AA.
XX
AC   AAE23980;
XX
DT   23-SEP-2002  (first entry)
XX
DE   Human LP220 secreted protein.
XX
KW   Human; secreted protein; atherosclerosis; Alzheimer's disease; LP220;
KW   diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
KW   rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
KW   reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
KW   gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;
KW   chromosome 11q13.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..16
FT                   /label= Signal_peptide
FT   Protein         17..635
FT                   /note= "Mature human LP220 secreted protein"
XX
PN   WO200226801-A2.
XX
PD   04-APR-2002.
XX
PF   14-SEP-2001; 2001WO-US26026.
XX
PR   28-SEP-2000; 2000US-236088P.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Su EW,  Wang H;
XX
DR   WPI; 2002-471259/50.
```

```
DR   N-PSDB; AAD38695.
XX
PT   Novel proteins and polynucleotides of secreted proteins useful for
PT   treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,
PT   diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury
PT   -
XX
PS   Claim 8; Page 127-129; 145pp; English.
XX
CC   The invention relates to human secreted polypeptides designated LP095,
CC   LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic
CC   acid molecules encoding such polypeptides. Novel secreted proteins of
CC   the invention are used for treating diseases such as atherosclerosis,
CC   Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,
CC   arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe
CC   combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,
CC   reperfusion injury, neoplasms and cancer especially liver cancer. They
CC   are also used for wound healing. Polynucleotides of the invention can
CC   be used to generate transgenic animals or knock out animals, which in
CC   turn, are useful in the development and screening of therapeutically
CC   useful reagents for use in the treatment of diseases associated with
CC   LP polypeptide associated activity. They are also used in gene therapy.
CC   The present sequence is human LP220 secreted protein. LP220 gene is
CC   located on chromosome 11q13.
XX
SQ   Sequence   635 AA;

  Query Match             40.7%;  Score 1667.5;  DB 23;  Length 635;
  Best Local Similarity   46.6%;  Pred. No. 1.7e-118;
  Matches  365;  Conservative   83;  Mismatches  177;  Indels  159;  Gaps   13;

Qy           8 LLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHI 67
               ||| | |     |||  |||||||||| |||  :|||||||::|||||||||  |||  :
Db           9 LLASGAA-----ACPLPCVCQNLSESLSTLCAHRGLLFVPPNVDRRTVELRLADNFIQAL 63

Qy          68 SRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQ 127
                  || ||||||||||||| |: |   :| ||||||||||| |||  ||  :||| ||||
Db          64 GPPDFRNMTGLVDLTLSRNAITRIGARAFGDLESLRSLHLDGNRLVELGTGSLRGPVNLQ 123

Qy         128 HLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDH 187
               |||:: |||| ||  ||:||| :|||||||||||  :||  :  |  || |:|||||:|
Db         124 HLILSGNQLGRIAPGAFDDFLESLEDLDLSYNNLRQVPWAGIGAMPALHTLNLDHNLIDA 183

Qy         188 IAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCN 247
               :  | || | :|:|||||||||  | |||:|:| : :   |:|   ||  || |||||||
Db         184 LPPGAFAQLGQLSRLDLTSNRLATLAPDPLFSRGRDA--EASP--APLVLSFSGNPLHCN 239

Qy         248 CELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAAT 307
               ||||||||| | |||||| || || ||||| | | || |||||| :|| :| ||||| ||
Db         240 CELLWLRRLARPDDLETCASPPGLAGRYFWAVPEGEFSCEPPLIARHTQRLWVLEGQRAT 299

Qy         308 LKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGE 367
               |:|:|:|||:| :||| |||||||||||   : ||||:| :| : |:| :|||| | |||
Db         300 LRCRALGDPAPTMHWVGPDDRLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATNPAGE 359

Qy         368 ATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVL 427
               ||| ||: :: |||  ||::    |    |||  |::|:  | |:   ||       ||
Db         360 ATARVELRVLALPHGGNSSAEGGRPGP--SDIAASARTAAEGEGTLESEP-------AVQ 410

Qy         428 VSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGY 487
               |:||| || || |   : |  | |:|:||| |:|| ||||::|||:  |:: :|| |  |
Db         411 VTEVTATSGLVSWGPGRPADPVWMFQIQYNSSEDETLIYRIVPASSHHFLLKHLVPGADY 470

Qy         488 DLCVLAMWDDTA-TTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATL 546
               |||:||:      : |||| ::||| | |    | | :: : :||||: : :||::|| |
Db         471 DLCLLALSPAAGPSDLTATRLLGCAHFSTLPASPLCHALQAHVLGGTLTVAVGGVLVAAL 530

Qy         547 LVFIVILMVRYK-VCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPPKVVVR 605
               ||||| |:|| :   |   | |:    |:| |||||     ||  |   ||: ||
Db         531 LVFIVALLVRGRGAGNGRLPLKL----SHVQSQTNGG----PSPTPKAHPPRSPP----- 577

Qy         606 NELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAFASLDL 665
                                                        |||
```

```
Db         578 -----------------------------------------PRP---------------- 580

Qy         666 KSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKRSHSFD 725
                                                                    :|| | |
Db         581 -----------------------------------------------------QRSCSLD 587

Qy         726 MGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSEWVM 785
               :||           | |   |::   | :|| ||:| ||        | ||  ||:| :
Db         588 LGD----------AGCYGYARRLGGAWARRSHSVHGGLLG------AGCRGVGGSAERLE 631

Qy         786 ESTV 789
               || |
Db         632 ESVV 635


RESULT 9
ABG13536
ID   ABG13536 standard; Protein; 322 AA.
XX
AC   ABG13536;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #13527.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS77723.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID NO 43895; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
```

```
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   322 AA;

  Query Match             40.4%;  Score 1656;  DB 22;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 5e-118;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        468 MIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHS 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHS 60

Qy        528 QILGGTMILVIGGIIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPP 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QILGGTMILVIGGIIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPP 120

Qy        588 SSAPAGAPPQGPPKVVVRNELLDFTASLARASDSSSSSLGSGEAAGLGRAPWRIPPSAP 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SSAPAGAPPQGPPKVVVRNELLDFTASLARASDSSSSSLGSGEAAGLGRAPWRIPPSAP 180

Qy        648 RPKPSLDRLMGAFASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARAR 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RPKPSLDRLMGAFASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARAR 240

Qy        708 SLLPLPLEGKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFE 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SLLPLPLEGKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFE 300

Qy        768 ESDLVGARGTFGSSEWVMESTV 789
              ||||||||||||||||||||||
Db        301 ESDLVGARGTFGSSEWVMESTV 322

RESULT 10
AAU32870
ID   AAU32870 standard; Protein; 636 AA.
XX
AC   AAU32870;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #3361.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US08656.
XX
PR   18-APR-2000; 2000US-0552929.
PR   26-JAN-2001; 2001US-0770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-611725/70.
XX
PT   Nucleic acids encoding a range of human polypeptides, useful in genetic
PT   vaccination, testing and therapy -
XX
PS   Claim 20; Page 682; 765pp; English.
XX
CC   The invention relates to novel human secreted polypeptides. The
CC   polypeptides and antibodies to the polypeptides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of polypeptide. The polypeptides are also useful for
CC   identifying agents (agonists and antagonists) that bind to them. Cells
CC   expressing the proteins are useful for identifying a therapeutic agent
CC   for use in treatment of a pathology related to aberrant expression or
CC   physiological interactions of the polypeptide. Vectors comprising
CC   the nucleic acids encoding the polypeptides and cells genetically
CC   engineered to express them are also useful for producing the proteins.
CC   The proteins are useful in genetic vaccination, testing and
CC   therapy, and can be used as nutritional supplements. They may be used to
CC   increase stem cell proliferation; to regulate haematopoiesis; and in
CC   bone, cartilage, tendon  and/or nerve tissue growth or regeneration;
CC   immune suppression and/or stimulation; as anti-inflammatory agents; and
CC   in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC   sequences of novel human secreted proteins of the invention.
XX
SQ   Sequence   636 AA;

  Query Match             38.8%;  Score 1586.5;  DB 22;  Length 636;
  Best Local Similarity   53.5%;  Pred. No. 2.6e-112;
  Matches  315;  Conservative   88;  Mismatches  171;  Indels   15;  Gaps    6;

Qy         16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
              |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db         23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy         76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
              |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db         83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy        136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
              |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |: 
Db        143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASVPAGAFSP 202

Qy        196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
              | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db        203 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 259

Qy        253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
              |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db        260 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 319

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db        320 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 379

Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSD-ITGSSKTSRGGGGSGGGEPPKSPPERAVLVS 429
              |:::     | |:||||   ||   | | :|  |  |     :  |     | :| | |:
Db        380 ELTVGPPPPPQLANSTSCDPPPGRGLPDALTPPSAASASAKVADTG----PPTDRGVQVT 435

Qy        430 EVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDL 489
              |   |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  |||
Db        436 EHGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDL 495

Qy        490 CVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVF 549
              ||||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::|||
Db        496 CVLAVYEDSATGLTATRPVGCARFSTEPSLRPCGAPHAPFLGGTMIIALGGVIVASVLVF 555

Qy        550 IVILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
              | :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db        556 IFVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 604

RESULT 11
AAB84469
ID   AAB84469 standard; Protein; 628 AA.
XX
AC   AAB84469;
XX
DT   05-SEP-2001  (first entry)
XX
DE   Amino acid sequence of an interferon omega-1 like protein NOV2.
XX
KW   Interferon omega-1 like protein; Nov2; membrane bound protein;
KW   secreted protein;. spermatogenesis; male infertility; neoplasia;
```

```
KW   blood circulation; immunological disorder; autoimmune disease;
KW   inflammatory disease; cardiovascular disease; metabolic disease;
KW   cancer; viral infection; acute lymphoblastic leukemia; glioma;
KW   neurological disease; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disorder; hematopoietic disorder; neurodegenerative disorder;
KW   immune disorder; hematopoietic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200142471-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US33463.
XX
PR   09-DEC-1999;   99US-0169887.
PR   10-DEC-1999;   99US-0170230.
PR   07-DEC-2000; 2000US-0170230.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Prayaga SK,  Shimkets RA;
XX
DR   WPI; 2001-381691/40.
DR   N-PSDB; AAH25568.
XX
PT   Novel polypeptides designated as NOV polypeptides, useful in detection,
PT   prevention and treatment of a broad range of pathological states  -
XX
PS   Claim 1; Page 16; 92pp; English.
XX
CC   The present sequence represents an interferon omega-1 like protein,
CC   designated NOV2. The protein is a membrane bound and secreted protein.
CC   NOV polypeptides and polynucleotides are useful for treating or
CC   preventing a NOV-associated disorder, and in the manufacture of a
CC   medicament for treating a syndrome associated with a human disease such
CC   as NOV-associated disorder. They are also useful for treating disorders
CC   or syndromes including those involved in development, differentiation
CC   and activation of thymic immune cells, pathologies related to
CC   spermatogenesis and male infertility, diagnosis of several human
CC   neoplasias, in disease or pathologies of cells in blood circulation
CC   such as red blood cells and platelets, various immunological disorders
CC   and/or pathologies, autoimmune and inflammatory diseases,
CC   cardiovascular diseases, metabolic diseases, cancer growth and
CC   metastasis, viral infections, cancer therapy, acute lymphoblastic
CC   leukemia, in gliomas, neurological diseases, neurodegenerative disorders,
CC   Alzheimer's disease, Parkinson's disorder, and hematopoietic disorders.
CC   NOV polypeptides and polynucleotides are also useful for treating
CC   neurodegenerative disorders, immune disorders and hematopoietic
CC   disorders.
XX
SQ   Sequence   628 AA;

  Query Match             38.4%;  Score 1573;  DB 22;  Length 628;
  Best Local Similarity   53.1%;  Pred. No. 2.7e-111;
  Matches  312;  Conservative   88;  Mismatches  174;  Indels   14;  Gaps    5;

Qy         16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
              |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db         23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy         76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
              |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db         83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy        136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
              |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db        143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASVPAGAFSR 202

Qy        196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
              | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db        203 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 259
```

```
Qy        253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
              |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db        260 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 319

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db        320 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 379

Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db        380 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 435

Qy        431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                 |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db        436 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 495

Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db        496 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 555

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||:  : |  :|: | ||:| ||||||  | |: ||    |
Db        556 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 603

RESULT 12
AAG67523
ID   AAG67523 standard; Protein; 628 AA.
XX
AC   AAG67523;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human secreted polypeptide.
XX
KW   Human; secreted polypeptide; nervous disease; muscular disease; tumour;
KW   gastrointestinal ulceration; spinal cord disease; trachea disease;
KW   thyroid gland disease; ovary disease; prostate disease; heart disease;
KW   renal gland disease; small intestine disease; thymus disease;
KW   lymph node disease; muscular system disease; colon disease;
KW   lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
KW   myocardial infarction; angioplasty; liver disease; coagulation disorder;
KW   microbial disease; immune disorder; inflammation; transplant rejection;
KW   bone thickness; bone density; ferroxidase loss; apoptosis;
KW   vascular smooth cell proliferation; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200166690-A2.
XX
PD   13-SEP-2001.
XX
PF   05-MAR-2001; 2001WO-US07143.
XX
PR   06-MAR-2000; 2000US-0187107.
PR   13-MAR-2000; 2000US-0188916.
PR   03-OCT-2000; 2000US-0236874.
PR   03-OCT-2000; 2000US-0237846.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Agarwal P,  Murdoch PR,  Rizvi SK,  Smith RF,  Xiang Z;
XX
DR   WPI; 2001-570768/64.
DR   N-PSDB; AAH78215.
XX
PT   Novel isolated secreted polypeptide useful for treating nervous and
PT   muscular diseases, gastrointestinal ulceration, coagulation and immune
PT   disorders, microbial diseases, inflammation and transplant rejection -
XX
PS   Claim 1; Page 83-85; 102pp; English.
XX
```

```
CC   The present sequence represents a human secreted polypeptide. The
CC   secreted polypeptides and polynucleotides are useful for treating
CC   nervous and muscular diseases, for inhibiting tumour formation and
CC   metastasis, for treating gastrointestinal ulceration, for preventing
CC   and treating diseases in spinal cord, thyroid gland, ovary, prostate,
CC   renal gland, small intestine, heart, trachea, thymus, lymph node,
CC   muscular system and colon, for treating lipase deficiency in cystic
CC   fibrosis and pancreatitis, for treating undesirable clot formation
CC   such as myocardial infarction, during angioplasty and all surgical
CC   procedures that require decreased blood clot formation, for treating
CC   liver diseases, coagulation disorders and microbial diseases, for
CC   treating immune disorders, for treating inflammation and transplant
CC   rejection, for enhancing bone thickness and increasing bone density,
CC   for reducing the loss of essential ferroxidases, for suppressing
CC   apoptosis, and for regulating vascular smooth cell proliferation. They
CC   may also be used as vaccines.
XX
SQ   Sequence   628 AA;

  Query Match             38.4%;  Score 1573;  DB 22;  Length 628;
  Best Local Similarity   53.1%;  Pred. No. 2.7e-111;
  Matches  312;  Conservative   88;  Mismatches  174;  Indels   14;  Gaps    5;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db.         23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASVPAGAFSR 202

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         203 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 259

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         260 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 319

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         320 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 379

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         380 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 435

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         436 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 495

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         496 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 555

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         556 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 603

RESULT 13
AAG65805
ID   AAG65805 standard; Protein; 628 AA.
XX
AC   AAG65805;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Human leucine-rich repeat (LRR) family member, 33395 polypeptide.
XX
KW   LRR; leucine rich repeat; 33395; cytostatic; anti-HIV; antidiabetic;
KW   antiarthritic; neuroprotective; dermatological; immunosuppresive;
KW   antiinflammatory; antiasthmatic; antiulcer; antianaemic; hepatotropic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          27..58
FT                   /note= "N-terminal LRR domain"
FT   Domain          60..83
FT                   /note= "LRR domain"
FT   Domain          84..107
FT                   /note= "LRR domain"
FT   Domain          108..131
FT                   /note= "LRR domain"
FT   Domain          132..155
FT                   /note= "LRR domain"
FT   Domain          157..180
FT                   /note= "LRR domain"
FT   Domain          181..204
FT                   /note= "LRR domain"
FT   Domain          205..228
FT                   /note= "LRR domain"
FT   Domain          249..294
FT                   /note= "C-terminal LRR domain"
FT   Domain          310..368
FT                   /note= "immunoglobulin domain"
FT   Domain          425..505
FT                   /note= "fibronectin type III domain"
FT   Domain          17..534
FT                   /note= "predicted extracellular domain"
FT   Domain          535..559
FT                   /note= "predicted transmembrane domain"
FT   Domain          560..628
FT                   /note= "predicted cytoplasmic domain"
XX
PN   WO200172827-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US09470.
XX
PR   24-MAR-2000; 2000US-191863P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Glucksmann MA;
XX
DR   WPI; 2001-626254/72.
DR   N-PSDB; AAI66985, AAI66986.
XX
PT   New polypeptide for preventing or treating disorders associated with
PT   cellular adhesion, proliferation or differentiation, comprises
PT   polypeptide 33395, a member of the leucine rich repeat protein family
PT   -
XX
PS   Claim 9; Fig 1; 133pp; English.
XX
CC   The invention provides an isolated nucleic acid encoding a polypeptide
CC   of the leucine rich repeat (LRR) family, designated 33395. The 33395
CC   polypeptide can be expressed by standard recombinant methodology. The
CC   33395 polynucleotides and polypeptide can be used to prevent or treat
CC   disorders associated with 33395 expression, for example those involving
CC   aberrant cellular adhesion, proliferation or differentiation. Specific
CC   examples include meloblastoma, juvenile AIDS, diabetes mellitus,
CC   rheumatoid arthritis, systemic lupus erythematosis, multiple sclerosis,
CC   Chron's disease, ulcerative colitis, asthma, anemia, and  chronic active
CC   hepatitis. The present sequence represents the human 33395 polypeptide.
XX
SQ   Sequence   628 AA;

  Query Match             38.3%;  Score 1566;  DB 22;  Length 628;
```

```
  Best Local Similarity   52.9%;  Pred. No. 9.2e-111;
  Matches  311;  Conservative   88;  Mismatches  175;  Indels   14;  Gaps    5;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASVPAGAFSR 202

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         203 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 259

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         260 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 319

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         320 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 379

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         380 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 435

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         436 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 495

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||::: :|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         496 VLAVYEGSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 555

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         556 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 603

RESULT 14
ABG34079
ID   ABG34079 standard; Protein; 627 AA.
XX
AC   ABG34079;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human Pro peptide #51.
XX
KW   Human; PRO; secreted protein; transmembrane protein;
KW   genetic disorder; tumour; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200224888-A2.
XX
PD   28-MAR-2002.
XX
PF   29-AUG-2001; 2001WO-US27099.
XX
PR   01-SEP-2000; 2000US-229896P.
PR   05-SEP-2000; 2000US-230621P.
PR   22-SEP-2000; 2000US-235147P.
PR   10-NOV-2000; 2000WO-US30873.
PR   12-JAN-2001; 2001US-261878P.
PR   16-JAN-2001; 2001US-261910P.
PR   16-JAN-2001; 2001US-261939P.
PR   16-JAN-2001; 2001US-262150P.
PR   25-JAN-2001; 2001US-264395P.
PR   02-FEB-2001; 2001US-266421P.
PR   09-FEB-2001; 2001US-267623P.
PR   28-FEB-2001; 2001WO-US06520.
PR   09-MAR-2001; 2001US-274399P.
PR   03-APR-2001; 2001US-280982P.
PR   04-APR-2001; 2001US-282129P.
PR   04-APR-2001; 2001US-282199P.
PR   09-MAY-2001; 2001US-290589P.
PR   25-MAY-2001; 2001WO-US17092.
PR   01-JUN-2001; 2001WO-US17800.
PR   20-JUN-2001; 2001WO-US19692.
PR   29-JUN-2001; 2001WO-US21066.
PR   09-JUL-2001; 2001WO-US21735.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Eaton DL,  Filvaroff E,  Goddard A,  Grimaldi JC;
PI   Gurney AL,  Smith V,  Stephan J,  Watanabe CK,  Wood WI,  Zhang Z;
PI   Fong S;
XX
DR   WPI; 2002-362426/39.
DR   N-PSDB; ABK70011.
XX
PT   New PRO polypeptides and polynucleotides encoding the polypeptides,
PT   useful in gene therapy, chromosome identification, tissue typing, or
PT   for genetic analysis of individuals with genetic disorders  -
XX
PS   Claim 11; Figure 102; 218pp; English.
XX
CC   This invention relates to the cDNA and protein sequences of novel
CC   secreted and transmembrane polypeptides PRO polypeptides. The
CC   invention also comprises a method for producing the proteins of the
CC   invention by recombinant means and antibodies specific for the protein
CC   of the invention. The antibody may be used for detecting the PRO
CC   proteins of the invention and may be used to modify their activity.
CC   polynucleotides may be used as hybridisation probes for a cDNA library
CC   to isolate the full-length PRO cDNA or to isolate other cDNAs, to
CC   construct hybridisation probes for mapping the gene which encodes that
CC   PRO and for genetic analysis of individuals with genetic disorders, in
CC   assays to identify other proteins or molecules involved in binding
CC   reaction, to generate transgenic animals or knock-out animals which in
CC   turn are useful in the development and screening of therapeutically
CC   useful reagents, for chromosome identification, and tissue typing. The
CC   PRO polypeptides are useful in gene therapy, and as molecular weight
CC   markers for protein electrophoresis purposes. The sequences may
CC   also be used to detect overexpression on PRO polypeptides in cancerous
CC   tumours and for screening for differentially expressed genes using
CC   microarray technology. The present sequence represents a human PRO
CC   protein of the invention.
XX
SQ   Sequence   627 AA;

  Query Match             38.2%;  Score 1563.5;  DB 23;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.4e-110;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258
```

```
Qy        253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
              |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db        259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db        319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db        379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy        431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                 |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db        435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db        495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db        555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 15
AAE17484
ID   AAE17484 standard; Protein; 551 AA.
XX
AC   AAE17484;
XX
DT   22-APR-2002  (first entry)
XX
DE   Human leucine-rich repeat-8 (ZLRR8) protein #2.
XX
KW   Human; leucine-rich repeat-8; ZLRR8; cytostatic; gene therapy; leukaemia;
KW   endometrial adenocarcinoma; renal cell; colon; prostrate; retinoblastoma;
KW   cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;
KW   mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;
KW   rhabdomyosarcoma; genitourinary tract; chromosome 11q13.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /label= Signal_peptide
FT   Protein         19..551
FT                   /label= Mature_Zllr8_protein
XX
PN   WO200202604-A2.
XX
PD   10-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-US20999.
XX
PR   30-JUN-2000; 2000US-215446P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Thayer EC,  Sheppard PO,  Presnell SR;
XX
DR   WPI; 2002-154725/20.
DR   N-PSDB; AAD28124.
XX
PT   New leucine-rich repeat proteins and polynucleotides, useful for
PT   diagnosing and treating disorders related to abnormal cell growth e.g.
PT   retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,
PT   kidney and lung tumors  -
XX
PS   Claim 10; Page 67-69; 82pp; English.
XX
CC   The invention relates to an isolated polypeptide comprising leucine-rich
CC   repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9.
```

```
CC   Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of
CC   proteins of interest from a host cell and to monitor the secretion of
CC   proteins in general from cells and tissues. The Zlrr DNA and proteins
CC   are useful in diagnosing and treating disorders related to abnormal cell
CC   growth, including retinoblastoma, renal cell adenocarcinoma, endometrial
CC   adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphotic leukaemia,
CC   kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,
CC   colon adenocarcinoma, genitourinary tract transitional cell tumours,
CC   rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas
CC   and prostate adenocarcinoma. Zlrr protein is useful for identifying
CC   agonists and antagonists of the polypeptide, for drug design, to screen
CC   for cell metabolism effecting receptors, for analysis of cell phenotype,
CC   and as animal feed supplement and cell culture components. Zlrr DNA is
CC   also useful in gene therapy. The present sequence is human Zlrr8 protein.
CC   Zlrr8 gene is located on chromosome 11q13.
XX
SQ   Sequence   551 AA;

  Query Match             37.8%;  Score 1545.5;  DB 23;  Length 551;
  Best Local Similarity   57.6%;  Pred. No. 2.8e-109;
  Matches  311;  Conservative   70;  Mismatches  140;  Indels   19;  Gaps    6;

Qy          8 LLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHI 67
              ||| | |     |||  |||||||||| |||  :|||||||::|||||||||  |||  :
Db          9 LLASGAA-----ACPLPCVCQNLSESLSTLCAHRGLLFVPPNVDRRTVELRLADNFIQAL 63

Qy         68 SRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQ 127
                 || ||||||||||||| |: |   :| ||||||||||| |||  ||  :||| ||||
Db         64 GPPDFRNMTGLVDLTLSRNAITRIGARAFGDLESLRSLHLDGNRLVELGTGSLRGPVNLQ 123

Qy        128 HLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDH 187
              |||:: |||| ||  ||:||| :|||||||||||  :||  :  |  || |:|||||:|
Db        124 HLILSGNQLGRIAPGAFDDFLESLEDLDLSYNNLRQVPWAGIGAMPALHTLNLDHNLIDA 183

Qy        188 IAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCN 247
              :  | || | :|:|||||||||  | |||:|:| : :   |:|   ||  || |||||||
Db        184 LPPGAFAQLGQLSRLDLTSNRLATLAPDPLFSRGRDA--EASP--APLVLSFSGNPLHCN 239

Qy        248 CELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAAT 307
              ||||||||| | |||||| || || ||||| | | || |||||| :|| :| ||||| ||
Db        240 CELLWLRRLARPDDLETCASPPGLAGRYFWAVPEGEFSCEPPLIARHTQRLWVLEGQRAT 299

Qy        308 LKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGE 367
              |:|:|:|||:| :||| |||||||||||   : ||||:| :| : |:| :|||| | |||
Db        300 LRCRALGDPAPTMHWVGPDDRLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATNPAGE 359

Qy        368 ATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVL 427
              ||| ||: :: |||  ||::    |    |||  |::|:  | |:   ||       ||
Db        360 ATARVELRVLALPHGGNSSAEGGRPGP--SDIAASARTAAEGEGTLESEP-------AVQ 410

Qy        428 VSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGY 487
              |:||| || || |   : |  | |:|:||| |:|| ||||::|||:  |:: :|| |  |
Db        411 VTEVTATSGLVSWGPGRPADPVWMFQIQYNSSEDETLIYRIVPASSHHFLLKHLVPGADY 470

Qy        488 DLCVLAMWDDTA-TTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATL 546
              |||:||:      : |||| ::||| | |    | | :: : :||||: : :||::|| |
Db        471 DLCLLALSPAAGPSDLTATRLLGCAHFSTLPASPLCHALQAHVLGGTLTVAVGGVLVAAL 530


Search completed: June  8, 2003, 21:28:18
Job time : 68 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          June  8, 2003, 21:27:43 ; Search time 36 Seconds
                                           (without alignments)
                                           644.852 Million cell updates/sec

Title:           US-09-831-846-2
Perfect score:   4094
Sequence:        1 METLLGGLLAFGMAFAVVDA..........DLVGARGTFGSSEWVMESTV 789

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:      262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                 1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                 2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                 3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                 4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                 5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                 6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                     %
Result             Query
  No.    Score     Match Length DB  ID                         Description
 ----------------------------------------------------------------------------
    1      368      9.0     708  4  US-09-131-648-2             Sequence 2, Appli
    2      368      9.0    1091  3  US-08-986-485-5             Sequence 5, Appli
    3      351      8.6    1101  3  US-08-986-485-2             Sequence 2, Appli
    4      299      7.3     673  4  US-09-063-950-2             Sequence 2, Appli
    5    284.5      6.9    1525  3  US-09-191-647-2             Sequence 2, Appli
    6    284.5      6.9    1525  4  US-09-540-245A-2            Sequence 2, Appli
    7    284.5      6.9    1525  4  US-09-540-153-2             Sequence 2, Appli
    8      268      6.5    1523  4  US-09-182-024A-2            Sequence 2, Appli
    9      263      6.4    1480  3  US-09-191-647-7             Sequence 7, Appli
   10      263      6.4    1480  4  US-09-540-245A-7            Sequence 7, Appli
   11      263      6.4    1480  4  US-09-540-153-7             Sequence 7, Appli
   12      263      6.4    1480  4  US-09-182-024A-5            Sequence 5, Appli
   13      263      6.4    1480  5  PCT-US91-09055-2            Sequence 2, Appli
   14    260.5      6.4     222  5  PCT-US91-09055-3            Sequence 3, Appli
   15      259      6.3     231  3  US-08-986-485-7             Sequence 7, Appli
   16      248      6.1     603  1  US-08-190-802A-50           Sequence 50, Appl
   17      248      6.1     603  4  US-08-477-346-50            Sequence 50, Appl
   18      248      6.1     603  4  US-08-473-089-50            Sequence 50, Appl
   19      248      6.1     603  4  US-08-487-072A-50           Sequence 50, Appl
   20    242.5      5.9     353  3  US-08-986-485-6             Sequence 6, Appli
   21    241.5      5.9     605  1  US-08-190-802A-49           Sequence 49, Appl
   22    241.5      5.9     605  4  US-08-477-346-49            Sequence 49, Appl
   23    241.5      5.9     605  4  US-08-473-089-49            Sequence 49, Appl
   24    241.5      5.9     605  4  US-08-487-072A-49           Sequence 49, Appl
   25    236.5      5.8     649  4  US-09-188-930-305           Sequence 305, App
   26    235.5      5.8     605  4  US-09-063-950-5             Sequence 5, Appli
   27      224      5.5     373  4  US-09-724-864-43            Sequence 43, Appl
   28      222      5.4     180  3  US-08-986-485-8             Sequence 8, Appli
   29      222      5.4     224  5  PCT-US91-09055-4            Sequence 4, Appli
   30      203      5.0     821  1  US-08-339-578-2             Sequence 2, Appli
   31      201      4.9     320  1  US-07-613-083B-1            Sequence 1, Appli
   32      198      4.8     610  1  US-07-821-717B-6            Sequence 6, Appli
   33      198      4.8     610  1  US-08-119-262B-6            Sequence 6, Appli
   34      198      4.8     610  1  US-08-135-929A-11           Sequence 11, Appl
   35      198      4.8     610  1  US-08-234-265A-11           Sequence 11, Appl
   36      197      4.8     560  3  US-08-592-500-2             Sequence 2, Appli
   37      197      4.8     560  3  US-08-195-006-2             Sequence 2, Appli
   38      197      4.8     560  4  US-09-063-950-4             Sequence 4, Appli
   39      197      4.8     560  5  PCT-US94-07644A-2           Sequence 2, Appli
   40      189      4.6     434  4  US-09-540-245A-19           Sequence 19, Appl
   41    186.5      4.6     375  1  US-08-303-238-2             Sequence 2, Appli
   42    186.5      4.6     375  4  US-08-458-834-2             Sequence 2, Appli
   43      185      4.5     612  2  US-08-359-705B-8            Sequence 8, Appli
   44      185      4.5     612  2  US-08-286-846A-8            Sequence 8, Appli
   45      185      4.5     612  2  US-08-457-880A-8            Sequence 8, Appli

                                   ALIGNMENTS

RESULT 1
US-09-131-648-2
; Sequence 2, Application US/09131648
; Patent No. 6168920
; GENERAL INFORMATION:
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Yue, Henry
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Patterson, Chandra
;  TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS
;  FILE REFERENCE: PF-0576 US
;  CURRENT APPLICATION NUMBER: US/09/131,648
;  CURRENT FILING DATE: 1998-08-10
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PERL Program
; SEQ ID NO 2
;   LENGTH: 708
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 2687731
US-09-131-648-2

  Query Match             9.0%;  Score 368;  DB 4;  Length 708;
  Best Local Similarity   22.6%;  Pred. No. 3.4e-22;
  Matches  154;  Conservative   98;  Mismatches  238;  Indels  192;  Gaps   27;

Qy          1 METLLGGLLAFGMAFAVVDA---CPKYCVCQ--------NLSESLGTL-CPSKGLLFVPP 48
              :  |||  ||       ||    ||: | |:        ::     |: |   |||  |
Db          8 IHVLLG--LAITTLVQAVDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLTFPA 65

Qy         49 DIDRRTVELRLGGNFIIHIS-RQDF-ANMTG----------------------------- 77
               :   |  | |  | |  |    ||  |:||
Db         66 RLPANTQILLLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLEE 125

Qy         78 ----------------LVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLR 121
                              | :| :: | :| | | :|: | :|  |||:||||  :
Db        126 NKLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFD 185

Qy        122 GLVNLQHLIVNNNQLGGIADEAF--------------------EDFLLTLEDLD-LSY-- 158
               | ||: |::  | :  | |  |                    :: |: ||:|: :|:
Db        186 ALPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYD 245

Qy        159 NNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFA------------------------ 194
              | |  :|  :::::|||  | |: | :: |  | |:
Db        246 NRLIKVPHVALQKVVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELISIDSLAV 305

Qy        195 -DLQKLARLDLTSN-RLQKLPPDPIFARSQASALTATPFA------------PPL-SFSF 239
               :|  | ::: |:| ||  : |:  |   :  :|     |            | |   |
```

```
Db        306 DNLPDLRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISI 365

Qy        240 GGNPLHCNCELLWLR------RLERDDDLETCGSPGGLKGRYFWHVREEEF-----VCEP 288
                ||: |:| : |:       |    | |  |  |   :|:   :||:  |     :| |
Db        366 HSNPIRCDCVIRWMNMNKTNIRFMEPDSL-FCVDPPEFQGQ---NVRQVHFRDMMEICLP 421

Qy        289 PLITQHTHKLLVLE-GQAATLKCKAIGDPSPLIHWVAPD-DRLVGN--SSRTAVYDNGTL 344
               :  :     | :| |   :  |:|  :| | |:|: |   :|: |  : :  |:  |||
Db        422 LIAPESFPSNLNVEAGSYVSFHCRATAEPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTL 481

Qy        345 DIFITTSQDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSK 404
              ||   | :: | :|||| |  |                       |  || :  :  ||
Db        482 DINGVTPKEGGLYTCIATNLVG-----------------------ADLKSVMIKVDGSFP 518

Qy        405 TSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVS----KSAPRVKMYQLQYNCSD 460
                  |  :             : : ::   | || |  |    ||: :   :       :
Db        519 QDNNGSLN-------------IKIRDIQANSVLVSWKASSKILKSSVKWTAF---VKTEN 562

Qy        461 DEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYP 520
                      ||:  | : : :|   | | :|:         |:   |   |    ||  :|
Db        563 SHAAQSARIPSDVKVYNLTHLNPSTEYKICI------DIPTIYQKNRKKCVNVTTKGLHP 616

Qy        521 QCQSMHSQILGGTMILVIGGII 542
                |  : :    |::  :||::
Db        617 D-QKEYEKNNTTTLMACLGGLL 637
```

```
RESULT 2
US-08-986-485-5
; Sequence 5, Application US/08986485
; Patent No. 6046030
;  GENERAL INFORMATION:
;    APPLICANT:  WU, SHUJIAN
;    APPLICANT:  SWEET, RAYMOND
;    APPLICANT:  TRUNEH, ALEMSEGED
;    TITLE OF INVENTION:  A HUMAN LIG-1 HOMOLOG (HLIG-1)
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/986,485
;      FILING DATE:  08-DEC-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/059,448
;      FILING DATE:  22-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-70264
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1091 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-986-485-5
```

```
  Query Match             9.0%;  Score 368;  DB 3;  Length 1091;
  Best Local Similarity  21.3%;  Pred. No. 6.6e-22;
  Matches  205;  Conservative   92;  Mismatches  326;  Indels  340;  Gaps   33;

Qy         30 LSESLGTLCPSKGLL-------FVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVDLT 82
              || || ||  ||  :       |  |    |  :| |  | |  |    |  :  |  |
Db        188 LSRSLLTLRLSKNRITQLPVKAFKLP----RLTQLDLNRNRIRLIEGLTFQGLDSLEVLR 243

Qy         83 LSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADE 142
              | || || :   :|  |  :  |||: | |  :   :| ||  |  | ::|| :  |  :
Db        244 LQRNNISRLTDGAFWGLSKMHVLHLEYNSLVEVNSGSLYGLTALHQLHLSNNSISRIQRD 303

Qy        143 AFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLARL 202
               :  |   | :| ||:|||  |  :|:  : :|  | | || : ||||| |  |: |  |
Db        304 GW-SFCQKLHELILSFNNLTRLDEESLAELSSLSILRLSHNAISHIAEGAFKGLKSLRVL 362

Qy        203 DLTSNRLQKLPPDPIFARSQASALT-ATPFAPPLS--------------FSFGGNPLH- 245
              ||  | :     |   | :    |:  | |   :                 : | | :
Db        363 DLDHNEISGTIEDTSGAFTGLDNLSKLTLFGNKIKSVAKRAFSGLESLEHLNLGENAIRS 422

Qy        246 -----------------------CNCELLWL-----RRLERDDDLETCGSPGGLKGRYFW 277
                                     |:|:| ||      |: :     ||  |  |||:  :
Db        423 VQFDAFAKMKNLKELYISSESFLCDCQLKWLPPWLMGRMLQAFVTATCAHPESLKGQSIF 482

Qy        278 HVREEEFVCE----PPLITQ---------------------------------------- 293
               |  : |||:    | :|||
Db        483 SVLPDSFVCDDFPKPQIITQPETTMAVVGKDIRFTCSAASSSSSPMTFAWKKDNEVLANA 542

Qy        294 ----------------------------------------------HTHK---------- 297
                                                            ::||
Db        543 DMENFAHVRAQDGEVMEYTTILHLRHVTFGHEGRYQCIITNHFGSTYSHKARLTVNVLPS 602

Qy        298 -------LLVLEGQAATLKCKAIGDPSPLIHW-----------------VAPDDRLVGNS 333
                     : :  |  |  |:| | | |:| | |                 | |||
Db        603 FTKIPHDIAIRTGTTARLECAATGHPNPQIAWQKDGGTDFPAARERRMHVMPDD------ 656

Qy        334 SRTAVYDNGTLDIFITTS---QDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTA 390
                         |:|  |     | | ::| | |:||  :|   ::::: | |       |
Db        657 -----------DVFFITDVKIDDMGVYSCTAQNSAGSVSANATLTVLETPSL-------A 698

Qy        391 PPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVK 450
               |                              :| | | |   | |    :     ||:
Db        699 VPLE----------------------------DRVVTVGE---TVAFQCKATGSPTPRIT 727

Qy        451 MYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTAT-TLTATNIVG 509
                :     |  |   :   |  |:  || |:            | ||    |   :| :|
Db        728 WLKGGRPLSLTE--RHHFTP-GNQLLVVQNV------------MIDDAGRYTCEMSNPLG 772

Qy        510 CAQFFTKADY---PQCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRYKVCNHEAPS 566
                :  ::      | |:      :|   | |:  |:: :|:   :|   | |
Db        773 TERAHSQLSILPTPGCRK-DGTTVGIFTIAVVCSIVLTSLVWVCIIYQTRKK-------- 823

Qy        567 KMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGP----PKVVVRNE--------------- 607
                    |  || ||  :   |   |:    ||      : ||| |
Db        824 ------SEEYSVTNTDETIVPPDVPSYLSSQGTLSDRQETVVRTEGGHQANGHIESNGVC 877

Qy        608 LLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPR-PKPSLDRLMGAF----AS 662
              | |   ||    |  |::        |  | ||::   | |   |      |:      |
Db        878 LRD--PSLFPEVDIHSTTCRQPKLCVGYTREPWKVTEKADRTAAPHTTAHSGSAVCSDCS 935

Qy        663 LDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKRSH 722
               |     :    ||  | |  :    | |  :  |               |  | |  ||
Db        936 TDTAYHPQPVPRDSGQP-GTASSQELRQHDREYSP-------------HHPYSGTADGSH 981

Qy        723 SFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSE 782
              :         : | : |  :                    :||  ::    | |  | |
Db        982 TL--------SGGSLYPSNHD--------------------RILPSLKNKAASADGN-GDSS 1014

Qy        783 WVM 785
              | :
Db       1015 WTL 1017
```

```
RESULT 3
US-08-986-485-2
; Sequence 2, Application US/08986485
; Patent No. 6046030
;  GENERAL INFORMATION:
;    APPLICANT:  WU, SHUJIAN
;    APPLICANT:  SWEET, RAYMOND
;    APPLICANT:  TRUNEH, ALEMSEGED
;    TITLE OF INVENTION:  A HUMAN LIG-1 HOMOLOG (HLIG-1)
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/986,485
;      FILING DATE:  08-DEC-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/059,448
;      FILING DATE:  22-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-70264
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1101 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-986-485-2

  Query Match             8.6%;  Score 351;  DB 3;  Length 1101;
  Best Local Similarity   21.6%;  Pred. No. 1.8e-20;
  Matches  168;  Conservative   77;  Mismatches  233;  Indels  298;  Gaps   23;

Qy          30 LSESLGTLCPSKGLL-------FVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVDLT 82
               || || ||  ||  :       |  |    |  :| |  | |  |    |  :  |  |
Db         193 LSRSLLTLRLSKNRITQLPVRAFKLP----RLTQLDLNRNRIRLIEGLTFQGLNSLEVLK 248

Qy          83 LSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADE 142
               | || || :   :|  |  :  |||: : |  :   :| ||  |  | ::|| :  |  :
Db         249 LQRNNISKLTDGAFWGLSKMHVLHLEYDSLVEVNSGSLYGLTALHQLHLSNNSIARIHRK 308

Qy         143 AFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLARL 202
                :  |   | :| ||:|||  |  |:|  : :|  | | || : ||||| |  |: |  |
Db         309 GW-SFCQKLHELVLSFNNLTRLDEESLAELSSLSVLRLSHNSISHIAEGAFKGLRSLRVL 367

Qy         203 DLTSNRLQKLPPDP--IFARSQASALTATPFAPPLS---------------FSFGGN--- 242
               ||  | :     |    |:  :      | |   :                 : |||
Db         368 DLDHNEISGTIEDTSGAFSGLEFGHSKLTLFGNKIKSVAKRAFSGLEGLEHLNLGGNAIR 427

Qy         243 --------------PLH-------CNCELLWL-----RRLERDDDLETCGSPGGLKGRYF 276
                              ||       |:|:| ||      |: :     ||  |  |||:
Db         428 SVQFDAFVKMKNLKELHISSDSFLCDCQLKWLPPWLIGRMLQAFVTATCAHPESLKGQSI 487

Qy         277 WHVREEEFVCE----PPLITQ--------------------------------------- 293
               : |  | |||:    | :|||
Db         488 FSVPPESFVCDDFLKPQIITQPETTMAMVGKDIRFTCSAASSSSSPMTFAWKKDNEVLTN 547
```

```
Qy         294 -----------------------------------------------HTHK--------- 297
                                                              ::||
Db         548 ADMENFVHVHAQDGEVMEYTTILHLRQVTFGHEGRYQCVITNHFGSTYSHKARLTVNVLP 607

Qy         298 --------LLVLEGQAATLKCKAIGDPSPLIHW-----------------VAPDDRLVGN 332
                       : :     | |:| | | |:| | |                 | |||
Db         608 SFTKTPHDITIRTTTVARLECAATGHPNPQIAWQKDGGTDFPAARERRMHVMPDD----- 662

Qy         333 SSRTAVYDNGTLDIFITTS---QDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRT 389
                           |:|  |     |:| ::| | |:||  :|   ::::: | |
Db         663 ------------DVFFITDVKIDDAGVYSCTAQNSAGSISANATLTVLETPSL------V 704

Qy         390 APPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPRV 449
                | : |                                |  |  | ||   :     ||:
Db         705 VPLEDR--------------------------------VVSVGETVALQCKATGNPPPRI 732

Qy         450 KMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVG 509
                 ::     |  |   : : |  |:  || |:|:            |    |   :| :|
Db         733 TWFKGDRPLSLTE--RHHLTP-DNQLLVVQNVVAE-----------DAGRYTCEMSNTLG 778

Qy         510 CAQFFTKADYPQCQSMHSQI--------------LGGTMILVIGGIIVATLLVFIVILMV 555
                            :  |||:              :|   | |:  |:: :|:   :|
Db         779 ------------TERAHSQLSVLPAAGCRKDGTTVGIFTIAVVSSIVLTSLVWVCIIYQT 826

Qy         556 RYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGP----PKVVVRNE 607
               | |              |  || ||  :   |   |:    ||      : ||| |
Db         827 RKK--------------SEEYSVTNTDETVVPPDVPSYLSSQGTLSDRQETVVRTE 868

RESULT 4
US-09-063-950-2
; Sequence 2, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION: THEREFOR
;  FILE REFERENCE: MEI-019
;  CURRENT APPLICATION NUMBER: US/09/063,950C
;  CURRENT FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 673
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-063-950-2

  Query Match             7.3%;  Score 299;  DB 4;  Length 673;
  Best Local Similarity   23.7%;  Pred. No. 1.8e-16;
  Matches  169;  Conservative   77;  Mismatches  268;  Indels  198;  Gaps   28;

Qy           8 LLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHI 67
               ||| |     |  ||  | |   |:     | ::    || |:   || | :  | |  :
Db          14 LLALGPG---VQGCPSGCQC---SQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGITML 67

Qy          68 SRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQ 127
                   || : ||  | ||:| |: :    |  | :| :| | :|||  :  :| |||  |:
Db          68 DAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLE 127

Qy         128 HLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNL--- 184
                | :  |::  |   || | |  | :| |  | |  ||     |:  |  | | ||
Db         128 RLYLGKNRIRHIQPGAF-DTLDRLLELKLQDNELRALP---PLRLPRLLLLDLSHNSLLA 183

Qy         185 --------------------LDHIAEGTFADLQKLARLDLTSNRLQKLPPD--------- 215
                                   |  : || |: |: |  ||:: |:|:::||
Db         184 LEPGILDTANVEALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRGLTR 243

Qy         216 --------------------PIFARSQASALTATPFAPPLSFSF--------GGNPLHCN 247
                                           | |:       ||  |          || :|
Db         244 LRLAGNTRIAQLRPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPFNCV 303
```

```
Qy        248 CELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCE-PPLITQHTHKLLVLEGQAA 306
              | | |     |:  : |  ||             ||  |  ||   ::  :|| ||   |
Db        304 CPLSWFGPWVRESHV-TLASP-------------EETRCHFPP---KNAGRLL-LELDYA 345

Qy        307 TLKCKA-------------IGDPSPLIH-----WVAPDDRLVGNSSRTAVYDNGTLDIFI 348
                 | |             : :|: |       |::|
Db        346 DFGCPATTTTATVPTTRPVVREPTALSSSLAPTWLSP----------------------- 382

Qy        349 TTSQDSGAFTCIAANA-AGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLS---------- 397
                       |  |  | :  :||   |  |  |       :  || : |:
Db        383 ---------TAPATEAPSPPSTAPPTVGPVPQP-------QDCPPSTCLNGGTCHLGTRH 426

Qy        398 --------DITGSSKTSRGGGGSGGGEPPKSP-PERAVL--VSEVTTTSALV--KWSVSK 444
                        ||    |: | |:     | :| | |::   :  |: ||  |  :  :
Db        427 HLACLCPEGFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGLQRYLQG 486

Qy        445 SAPRVKMYQLQYN--CSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTL 502
              |:  :::  :| |      |: |:   :|||   : |  |     | :||: :
Db        487 SSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCVMPL----GPGR 542

Qy        503 TATNIVGCAQFFTKADYPQCQSMH---SQILGGTMILVIGGIIVATLLVFIVILMVRYKV 559
                     | :  |    |   | |   :|   | : |:|   : | ||  :  :   | |
Db        543 VPEGEEACGEAHTP---PAVHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCV 599

Qy        560 CNHEAPSKMAAAVSNVYSQTNGAQP-----------PPPSSAPAG--APPQG 598
                  |   ||||  :      || |           | | :   |  | | |
Db        600 RRGRA---MAAAAQDKGQVGPGAGPLELEGVKVPLEPGPKATEGGGEALPSG 648
```

```
RESULT 5
US-09-191-647-2
; Sequence 2, Application US/09191647
; Patent No. 6046015
; GENERAL INFORMATION:
;  APPLICANT: Goodman, Corey
;  APPLICANT: Kid, Thomas
;  APPLICANT: Brose, Katja
;  APPLICANT: Tessier-Lavigne, Marc
;  TITLE OF INVENTION: Modulating Robo: Ligand Interactions
;  FILE REFERENCE: B98-031-3
;  CURRENT APPLICATION NUMBER: US/09/191,647
;  CURRENT FILING DATE: 1998-11-13
;  EARLIER APPLICATION NUMBER: 60/065,544
;  EARLIER FILING DATE: 1997-11-14
;  EARLIER APPLICATION NUMBER: 60/081,057
;  EARLIER FILING DATE: 1998-04-07
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1525
;   TYPE: PRT
;   ORGANISM: human
US-09-191-647-2

  Query Match             6.9%;  Score 284.5;  DB 3;  Length 1525;
  Best Local Similarity   29.2%;  Pred. No. 1.1e-14;
  Matches   84;  Conservative   36;  Mismatches  117;  Indels   51;  Gaps    6;

Qy          9 LAFGMAFAVVD-----ACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNF 63
              |:  |:  |:::     |||  | |   :      |    |  || :| | |  | | ||
Db         11 LSLGLVLAILNKVAPQACPAQCSCSGSTVD----CHGLALRSVPRNIPRNTERLDLNGNN 66

Qy         64 IIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGL 123
              |  |:: ||| :  |  | |  | || |:  :| ||: |  | |: | |    |    |
Db         67 ITRITKTDFAGLRHLRVLQLMENKISTIERGAFQDLKELERLRLNRNHLQLFPELLFLGT 126

Qy        124 VNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHN 183
                |                           |||| | :  :|  : |  |::  | ||:|
Db        127 AKLYR-------------------------LDLSENQIQAIPRKAFRGAVDIKNLQLDYN 161

Qy        184 LLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPL-SFSFGGN 242
               :  | :| |  |: |  | | :| : :|            :: :    | | :|    |
Db        162 QISCIEDGAFRALRDLEVLTLNNNNITRL------------SVASFNHMPKLRTFRLHSN 209
```

```
Qy        243 PLHCNCELL----WLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVC 286
               |:|:| |     |||:  |      |  |  |:|     |:: ||||
Db        210 NLYCDCHLAWLSDWLRKRPRVGLYTQCMGPSHLRGHNVAEVQKREFVC 257
```

```
RESULT 6
US-09-540-245A-2
; Sequence 2, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
;  APPLICANT:  Goodman, Corey
;  APPLICANT:  Kid, Thomas
;  APPLICANT:  Brose, Katja
;  APPLICANT:  Tessier-Lavigne, Marc
;  TITLE OF INVENTION:  Modulating Robo: Ligand Interactions
;  FILE REFERENCE:  B98-031-3
;  CURRENT APPLICATION NUMBER: US/09/540,245A
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER:  60/065,544
;  PRIOR FILING DATE:  1997-11-14
;  PRIOR APPLICATION NUMBER:  60/081,057
;  PRIOR FILING DATE:  1998-04-07
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1525
;   TYPE: PRT
;   ORGANISM: human
US-09-540-245A-2

  Query Match             6.9%;  Score 284.5;  DB 4;  Length 1525;
  Best Local Similarity   29.2%;  Pred. No. 1.1e-14;
  Matches   84;  Conservative   36;  Mismatches  117;  Indels   51;  Gaps    6;

Qy          9 LAFGMAFAVVD-----ACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNF 63
              |:  |:  |:::     |||  | |   :      |    |  || :| | |  | | ||
Db         11 LSLGLVLAILNKVAPQACPAQCSCSGSTVD----CHGLALRSVPRNIPRNTERLDLNGNN 66

Qy         64 IIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGL 123
              |  |:: ||| :  |  | |  | || |:  :| ||: |  | |: | |    |    |
Db         67 ITRITKTDFAGLRHLRVLQLMENKISTIERGAFQDLKELERLRLNRNHLQLFPELLFLGT 126

Qy        124 VNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHN 183
                |                           |||| | :  :|  : |  |::  | ||:|
Db        127 AKLYR-------------------------LDLSENQIQAIPRKAFRGAVDIKNLQLDYN 161

Qy        184 LLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPL-SFSFGGN 242
               :  | :| |  |: |  | | :| : :|            :: :    | | :|    |
Db        162 QISCIEDGAFRALRDLEVLTLNNNNITRL------------SVASFNHMPKLRTFRLHSN 209

Qy        243 PLHCNCELL----WLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVC 286
               |:|:| |     |||:  |      |  |  |:|     |:: ||||
Db        210 NLYCDCHLAWLSDWLRKRPRVGLYTQCMGPSHLRGHNVAEVQKREFVC 257
```

```
RESULT 7
US-09-540-153-2
; Sequence 2, Application US/09540153
; Patent No. 6270995
; GENERAL INFORMATION:
;  APPLICANT:  Goodman, Corey
;  APPLICANT:  Kid, Thomas
;  APPLICANT:  Brose, Katja
;  APPLICANT:  Tessier-Lavigne, Marc
;  TITLE OF INVENTION:  Modulating Robo: Ligand Interactions
;  FILE REFERENCE:  B98-031-3
;  CURRENT APPLICATION NUMBER: US/09/540,153
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER:  09/191,647
;  PRIOR FILING DATE:  1998-11-13
;  PRIOR APPLICATION NUMBER:  60/081,057
;  PRIOR FILING DATE:  1998-04-07
```

```
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1525
;   TYPE: PRT
;   ORGANISM: human
US-09-540-153-2

  Query Match              6.9%;  Score 284.5;  DB 4;  Length 1525;
  Best Local Similarity   29.2%;  Pred. No. 1.1e-14;
  Matches   84;  Conservative   36;  Mismatches  117;  Indels   51;  Gaps    6;

Qy          9 LAFGMAFAVVD-----ACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNF 63
              |: |:  |:::     |||  | |   :      |    |  || :| | |  | | ||
Db         11 LSLGLVLAILNKVAPQACPAQCSCSGSTVD----CHGLALRSVPRNIPRNTERLDLNGNN 66

Qy         64 IIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGL 123
              |  |:: ||| :  |  | |  | || |:  :| ||: |  | |: | |    |    |
Db         67 ITRITKTDFAGLRHLRVLQLMENKISTIERGAFQDLKELERLRLNRNHLQLFPELLFLGT 126

Qy        124 VNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHN 183
                |                           |||| | :  :|  : |  |::  | ||:|
Db        127 AKLYR-------------------------LDLSENQIQAIPRKAFRGAVDIKNLQLDYN 161

Qy        184 LLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPL-SFSFGGN 242
               :  | :| |  |: |  | | :| :  |            :: :    | | :|    |
Db        162 QISCIEDGAFRALRDLEVLTLNNNNITRL------------SVASFNHMPKLRTFRLHSN 209

Qy        243 PLHCNCELL----WLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVC 286
               |:|:| |     |||:  |      |  |  |:|     |:: ||||
Db        210 NLYCDCHLAWLSDWLRKRPRVGLYTQCMGPSHLRGHNVAEVQKREFVC 257

RESULT 8
US-09-182-024A-2
; Sequence 2, Application US/09182024A
; Patent No. 6342370
; GENERAL INFORMATION:
;  APPLICANT:  Connolly, Timothy
;  APPLICANT:  Rajput, Bhanu
;  TITLE OF INVENTION:  Human Slit Polypeptide and Polynucleotides Encoding
;  TITLE OF INVENTION:  Same
;  FILE REFERENCE:  640100-271
;  CURRENT APPLICATION NUMBER: US/09/182,024A
;  CURRENT FILING DATE:  1998-10-29
;  PRIOR APPLICATION NUMBER:  60/063,946
;  PRIOR FILING DATE:  1997-10-31
;  PRIOR APPLICATION NUMBER:  60/096,420
;  PRIOR FILING DATE:  1998-08-13
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1523
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-182-024A-2

  Query Match              6.5%;  Score 268;  DB 4;  Length 1523;
  Best Local Similarity   26.8%;  Pred. No. 2.5e-13;
  Matches   86;  Conservative   41;  Mismatches  130;  Indels   64;  Gaps    7;

Qy          9 LAFGMAFAVV------DACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGN 62
              ||  :| | |       |||  | |   |      |   ||  ||  | |    | |  |
Db         16 LALALALASVLSGPPAVACPTKCTCSAASVD----CHGLGLRAVPRGIPRNAERLDLDRN 71

Qy         63 FIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRG 122
               |  |:: ||| :  |  | |  | :| |:  :| ||: |  | |: |:|  | |   :
Db         72 NITRITKMDFAGLKNLRVLHLEDNQVSVIERGAFQDLKQLERLRLNKNKLQVLPELLFQS 131

Qy        123 LVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDH 182
                 |                           |||| | : |:|  : | : ::  | ||:
Db        132 TPKLTR-------------------------LDLSENQIQGIPRKAFRGITDVKNLQLDN 166
```

```
Qy        183 NLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPL-SFSFGG 241
              | :  | :| |  |: |  | | :| : ::             :|:    | : :
Db        167 NHISCIEDGAFRALRDLEILTLNNNNISRI------------LVTSFNHMPKIRTLRLHS 214

Qy        242 NPLHCNCELL----WLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHK 297
              | |:|:| |     |||:         | :|  |:|     |:::|:||  |
Db        215 NHLYCDCHLAWLSDWLRQRRTVGQFTLCMAPVHLRGFNVADVQKKEYVCPAP-------- 266

Qy        298 LLVLEGQAATLKCKAIGDPSP 318
                    :  :    :|  |||
Db        267 ----HSEPPSCNANSISCPSP 283

RESULT 9
US-09-191-647-7
; Sequence 7, Application US/09191647
; Patent No. 6046015
; GENERAL INFORMATION:
;  APPLICANT: Goodman, Corey
;  APPLICANT: Kid, Thomas
;  APPLICANT: Brose, Katja
;  APPLICANT: Tessier-Lavigne, Marc
;  TITLE OF INVENTION: Modulating Robo: Ligand Interactions
;  FILE REFERENCE: B98-031-3
;  CURRENT APPLICATION NUMBER: US/09/191,647
;  CURRENT FILING DATE: 1998-11-13
;  EARLIER APPLICATION NUMBER: 60/065,544
;  EARLIER FILING DATE: 1997-11-14
;  EARLIER APPLICATION NUMBER: 60/081,057
;  EARLIER FILING DATE: 1998-04-07
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 1480
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-191-647-7

  Query Match              6.4%;  Score 263;  DB 3;  Length 1480;
  Best Local Similarity   28.0%;  Pred. No. 6.3e-13;
  Matches   89;  Conservative   39;  Mismatches  102;  Indels   88;  Gaps   11;

Qy          6 GGLLAFGMAF------AVVDA-CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELR 58
              ||| : |:         : :| ||: | |  |:      |  :||  ||  |      |
Db         51 GGLGSVGIHIPGGGVGVITEARCPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106

Qy         59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
              | || :  |   ||  :| |  | |: | |  |:  || || ||  | : :| : ::|
Db        107 LQGNNLTVIYETDFQRLTKLRMLQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRR 166

Qy        119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                :|  :|: | ::|||:  : : ||:  |: || | |: |||  ||
Db        167 VFKGAQSLRSLQLDNNQITCLDEHAFKG-LVELEILTLNNNNLTSLP------------- 212

Qy        179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
                 ||:        |  | :|  | |: |                      |||
Db        213 ---HNI--------FGGLGRLRALRLSDN----------------------PFA------ 233

Qy        239 FGGNPLHCNCELLWLRRLERD----DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                     |:| | || |  |          | ||  |||:    : ::|| |    :|:|
Db        234 -------CDCHLSWLSRFLRSATRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 284

Qy        295 THKLLVLEGQAATLKCKA 312
                         | ::| |
Db        285 -----------APMECGA 291

RESULT 10
US-09-540-245A-7
; Sequence 7, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
;  APPLICANT:  Goodman, Corey
```

```
;  APPLICANT:  Kid, Thomas
;  APPLICANT:  Brose, Katja
;  APPLICANT:  Tessier-Lavigne, Marc
;  TITLE OF INVENTION:  Modulating Robo: Ligand Interactions
;  FILE REFERENCE:  B98-031-3
;  CURRENT APPLICATION NUMBER: US/09/540,245A
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER:  60/065,544
;  PRIOR FILING DATE:  1997-11-14
;  PRIOR APPLICATION NUMBER:  60/081,057
;  PRIOR FILING DATE:  1998-04-07
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 1480
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-540-245A-7

  Query Match             6.4%;  Score 263;  DB 4;  Length 1480;
  Best Local Similarity   28.0%;  Pred. No. 6.3e-13;
  Matches   89;  Conservative   39;  Mismatches  102;  Indels   88;  Gaps   11;

Qy          6 GGLLAFGMAF------AVVDA-CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELR 58
              ||| : |:         : :| ||: | |  |:      |  :||  ||  |      |
Db         51 GGLGSVGIHIPGGGVGVITEARCPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106

Qy         59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
              | ||  :  |   ||  :| |  | |: | |  |:  || || ||  |  : :| : ::|
Db        107 LQGNNLTVIYETDFQRLTKLRMLQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRR 166

Qy        119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                :|  :|: | ::|||:  : : ||:  |: || | |: |||  ||
Db        167 VFKGAQSLRSLQLDNNQITCLDEHAFKG-LVELEILTLNNNNLTSLP------------ 212

Qy        179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
                 ||:        |  | :|  | |: |                      |||
Db        213 ---HNI--------FGGLGRLRALRLSDN----------------------PFA------ 233

Qy        239 FGGNPLHCNCELLWLRRLERD----DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                     |:| | || |  |          | ||  |||:     : :|| |    :|:|
Db        234 -------CDCHLSWLSRFLRSATRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 284

Qy        295 THKLLVLEGQAATLKCKA 312
                         | ::| |
Db        285 -----------APMECGA 291


RESULT 11
US-09-540-153-7
; Sequence 7, Application US/09540153
; Patent No. 6270995
; GENERAL INFORMATION:
;  APPLICANT:  Goodman, Corey
;  APPLICANT:  Kid, Thomas
;  APPLICANT:  Brose, Katja
;  APPLICANT:  Tessier-Lavigne, Marc
;  TITLE OF INVENTION:  Modulating Robo: Ligand Interactions
;  FILE REFERENCE:  B98-031-3
;  CURRENT APPLICATION NUMBER: US/09/540,153
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER:  09/191,647
;  PRIOR FILING DATE:  1998-11-13
;  PRIOR APPLICATION NUMBER:  60/081,057
;  PRIOR FILING DATE:  1998-04-07
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 1480
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-540-153-7
```

```
  Query Match             6.4%;  Score 263;  DB 4;  Length 1480;
  Best Local Similarity   28.0%;  Pred. No. 6.3e-13;
  Matches   89;  Conservative   39;  Mismatches  102;  Indels   88;  Gaps   11;

Qy          6 GGLLAFGMAF------AVVDA-CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELR 58
              ||| : |:         : :| ||: | |  |:      |  :||  ||  |      |
Db         51 GGLGSVGIHIPGGGVGVITEARCPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106

Qy         59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
              | ||  :  |   ||  :| |  | |: | |  |:  || || ||  |  : :| : ::|
Db        107 LQGNNLTVIYETDFQRLTKLRMLQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRR 166

Qy        119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                :|  :|: | ::|||:  : : ||:  |: || | |: |||  ||
Db        167 VFKGAQSLRSLQLDNNQITCLDEHAFKG-LVELEILTLNNNNLTSLP------------ 212

Qy        179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
                 ||:        |  | :|  | |: |                      |||
Db        213 ---HNI--------FGGLGRLRALRLSDN----------------------PFA------ 233

Qy        239 FGGNPLHCNCELLWLRRLERD----DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                     |:| | || |  |          | ||  |||:     : :|| |    :|:|
Db        234 -------CDCHLSWLSRFLRSATRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 284

Qy        295 THKLLVLEGQAATLKCKA 312
                         | ::| |
Db        285 -----------APMECGA 291


RESULT 12
US-09-182-024A-5
; Sequence 5, Application US/09182024A
; Patent No. 6342370
; GENERAL INFORMATION:
;  APPLICANT:  Connolly, Timothy
;  APPLICANT:  Rajput, Bhanu
;  TITLE OF INVENTION:  Human Slit Polypeptide and Polynucleotides Encoding
;  TITLE OF INVENTION:  Same
;  FILE REFERENCE:  640100-271
;  CURRENT APPLICATION NUMBER: US/09/182,024A
;  CURRENT FILING DATE:  1998-10-29
;  PRIOR APPLICATION NUMBER:  60/063,946
;  PRIOR FILING DATE:  1997-10-31
;  PRIOR APPLICATION NUMBER:  60/096,420
;  PRIOR FILING DATE:  1998-08-13
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1480
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-182-024A-5

  Query Match             6.4%;  Score 263;  DB 4;  Length 1480;
  Best Local Similarity   28.0%;  Pred. No. 6.3e-13;
  Matches   89;  Conservative   39;  Mismatches  102;  Indels   88;  Gaps   11;

Qy          6 GGLLAFGMAF------AVVDA-CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELR 58
              ||| : |:         : :| ||: | |  |:      |  :||  ||  |      |
Db         51 GGLGSVGIHIPGGGVGVITEARCPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106

Qy         59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
              | ||  :  |   ||  :| |  | |: | |  |:  || || ||  |  : :| : ::|
Db        107 LQGNNLTVIYETDFQRLTKLRMLQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRR 166

Qy        119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                :|  :|: | ::|||:  : : ||:  |: || | |: |||  ||
Db        167 VFKGAQSLRSLQLDNNQITCLDEHAFKG-LVELEILTLNNNNLTSLP------------ 212

Qy        179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
                 ||:        |  | :|  | |: |                      |||
Db        213 ---HNI--------FGGLGRLRALRLSDN----------------------PFA------ 233
```

```
Qy        239 FGGNPLHCNCELLWLRRLERD----DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                     |:| | || |  |          | ||  |||:    : ::|| |    :|:|
Db        234 -------CDCHLSWLSRFLRSATRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 284

Qy        295 THKLLVLEGQAATLKCKA 312
                         | ::| |
Db        285 -----------APMECGA 291

RESULT 13
PCT-US91-09055-2
; Sequence 2, Application PC/TUS9109055
;  GENERAL INFORMATION:
;    APPLICANT:  Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
;    TITLE OF INVENTION:  Purified SLIT protein and Sequence Elements Thereof
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Yale University
;      ADDRESSEE:  Office of Cooperative Research
;      STREET:  246 Church Street
;      STREET:  Suite 401
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06510
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch. 800 Kb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh 6.0.5
;      SOFTWARE:  Microsoft Word 4.O
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/09055
;      FILING DATE:  19911127
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/624,135
;      FILING DATE:  7-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barth, Richard J.
;      REGISTRATION NUMBER:  28,180
;      REFERENCE/DOCKET NUMBER:  900964/RSB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 972-1400
;      TELEFAX:  (212) 370-1622
;      TELEX:  236268
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1480 amino acids
;      TYPE:  AMINO ACIDS
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  signal sequence
;      LOCATION:  1 to 36
;      IDENTIFICATION METHOD:  similarity to other signal
;      OTHER INFORMATION:  Directs Export
;      NAME/KEY:  Four Flank-LRR-Flank domains
;      LOCATION:  37 to 910
;      IDENTIFICATION METHOD:  Array of Flank-LRR-Flank
;      OTHER INFORMATION:  mediates adhesive events
;      NAME/KEY:  Tandem EGF-like repeats
;      LOCATION:  911 to 1150
;      IDENTIFICATION METHOD:  similarity to tandem EGF-like                    r
;      OTHER INFORMATION:  protein-protein interactions
;      NAME/KEY:  7th EGF-like repeat
;      LOCATION:  1353 to 1393
;      IDENTIFICATION METHOD:  similarity to epidermal growth                   f
;      OTHER INFORMATION:  Involvement in receptor-ligand
;      NAME/KEY:  Alternative splice segment
;      LOCATION:  1394 to 1404
;      IDENTIFICATION METHOD:  experimental
;      OTHER INFORMATION:  developmentally regulated
;      NAME/KEY:  COOH-terminal region
```

```
;      LOCATION:  1405 to 1480
;      IDENTIFICATION METHOD:  experimental
PCT-US91-09055-2

  Query Match             6.4%;  Score 263;  DB 5;  Length 1480;
  Best Local Similarity   28.0%;  Pred. No. 6.3e-13;
  Matches   89;  Conservative   39;  Mismatches  102;  Indels   88;  Gaps   11;

Qy          6 GGLLAFGMAF------AVVDA-CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELR 58
              ||| : |:         : :| ||: | |  |:      |  :||  ||  |      |
Db         51 GGLGSVGIHIPGGGVGVITEARCPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106

Qy         59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
              | || :  |   ||  :| |  | |: | |  |:  || || ||  | : :| : ::|
Db        107 LQGNNLTVIYETDFQRLTKLRMLQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRR 166

Qy        119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                :|  :|: | ::|||:  : : ||:  |: || | |: |||  ||
Db        167 VFKGAQSLRSLQLDNNQITCLDEHAFKG-LVELEILTLNNNNLTSLP------------- 212

Qy        179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
                 ||:        |  | :|  | |: |                      |||
Db        213 ---HNI--------FGGLGRLRALRLSDN----------------------PFA------ 233

Qy        239 FGGNPLHCNCELLWLRRLERD----DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                     |:| | || |  |          | ||  |||:    : ::|| |    :|:|
Db        234 -------CDCHLSWLSRFLRSATRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 284

Qy        295 THKLLVLEGQAATLKCKA 312
                         | ::| |
Db        285 -----------APMECGA 291

RESULT 14
PCT-US91-09055-3
; Sequence 3, Application PC/TUS9109055
;  GENERAL INFORMATION:
;    APPLICANT:  Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
;    TITLE OF INVENTION:  Purified SLIT protein and Sequence Elements Thereof
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Yale University
;      ADDRESSEE:  Office of Cooperative Research
;      STREET:  246 Church Street
;      STREET:  Suite 401
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06510
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch. 800 Kb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh 6.0.5
;      SOFTWARE:  Microsoft Word 4.O
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/09055
;      FILING DATE:  19911127
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/624,135
;      FILING DATE:  7-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barth, Richard J.
;      REGISTRATION NUMBER:  28,180
;      REFERENCE/DOCKET NUMBER:  900964/RSB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 972-1400
;      TELEFAX:  (212) 370-1622
;      TELEX:  236268
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  222 amino acids
;      TYPE:  AMINO ACIDS
```

```
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Flank-LRR-Flank 1
;      LOCATION:  1 to 222
;      IDENTIFICATION METHOD:  similarity to other Flank-LRR-                F
;      OTHER INFORMATION:  mediates adhesive events
PCT-US91-09055-3

  Query Match             6.4%;  Score 260.5;  DB 5;  Length 222;
  Best Local Similarity   28.4%;  Pred. No. 5.6e-14;
  Matches   84;  Conservative   35;  Mismatches   96;  Indels   81;  Gaps    9;

Qy          21 CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
               ||: | |  |:      |  :||  ||  |      | | ||  : |   ||  :| |
Db           1 CPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLELQGNNLTVIYETDFQRLTKLRM 56

Qy          81 LTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIA 140
               | |: | |  |:  || || ||  | : :| : ::|    :|  :|: | ::|||:  :
Db          57 LQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRRVFKGAQSLRSLQLDNNQITCLD 116

Qy         141 DEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLA 200
               : ||:  |: || | |: |||  ||                ||:        |  | :|
Db         117 EHAFKG-LVELEILTLNNNNLTSLP----------------HNI--------FGGLGRLR 151

Qy         201 RLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLERD- 259
                | |: |                      |||             |:| | || |  |
Db         152 ALRLSDN----------------------PFA-------------CDCHLSWLSRFLRSA 176

Qy         260 ---DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
                       | ||  |||:    : ::|| |    :|:|           | ::| |
Db         177 TRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH-----------APMECGA 219


RESULT 15
US-08-986-485-7
; Sequence 7, Application US/08986485
; Patent No. 6046030
;  GENERAL INFORMATION:
;    APPLICANT:  WU, SHUJIAN
;    APPLICANT:  SWEET, RAYMOND
;    APPLICANT:  TRUNEH, ALEMSEGED
;    TITLE OF INVENTION:  A HUMAN LIG-1 HOMOLOG (HLIG-1)
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/986,485
;      FILING DATE:  08-DEC-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/059,448
;      FILING DATE:  22-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-70264
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;
;      LENGTH:  231 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-986-485-7

  Query Match             6.3%;  Score 259;  DB 3;  Length 231;
  Best Local Similarity   29.1%;  Pred. No. 7.9e-14;
  Matches   81;  Conservative   33;  Mismatches   94;  Indels   70;  Gaps    8;

Qy          21 CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
               ||: | |  |:      |  :||  ||  |      | | ||  : |   ||  :| |
Db          17 CPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLELQGNNLTVIYETDFQRLTKLRM 72

Qy          81 LTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIA 140
               | |: | |  |:  || || ||  | : :| : ::|    :|  :|: | ::|||:  :
Db          73 LQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRRVFKGAQSLRSLQLDNNQITCLD 132

Qy         141 DEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLA 200
               : ||:  |: || | |: |||  ||                ||:        |  | :|
Db         133 EHAFKG-LVELEILTLNNNNLTSLP----------------HNI--------FGGLGRLR 167

Qy         201 RLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLERD- 259
                | |: |                      |||             |:| | || |  |
Db         168 ALRLSDN----------------------PFA-------------CDCHLSWLSRFLRSA 192

Qy         260 ---DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                       | ||  |||:    : ::|| |    :|:|
Db         193 TRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 228


Search completed: June 8, 2003, 21:32:09
Job time : 41 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         June  8, 2003, 21:30:28 ; Search time 55 Seconds
                                          (without alignments)
                                          1481.029 Million cell updates/sec

Title:          US-09-831-846-2
Perfect score:  4094
Sequence:       1 METLLGGLLAFGMAFAVVDA..........DLVGARGTFGSSEWVMESTV 789

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters:      392085

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                1:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                10:  /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                12:  /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
                13:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                14:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                    %
Result            Query
   No.    Score   Match Length DB  ID                          Description
------------------------------------------------------------------------------
     1     1566    38.3    628 10  US-09-815-626-2              Sequence 2, Appli
     2   1563.5    38.2    627  9  US-10-245-103-100           Sequence 100, App
     3   1563.5    38.2    627  9  US-10-245-107-100           Sequence 100, App
     4   1563.5    38.2    627  9  US-10-245-143-100           Sequence 100, App
     5   1563.5    38.2    627  9  US-10-245-771-100           Sequence 100, App
     6   1563.5    38.2    627  9  US-10-245-851-100           Sequence 100, App
     7   1563.5    38.2    627  9  US-10-245-883-100           Sequence 100, App
     8   1563.5    38.2    627  9  US-10-237-535-100           Sequence 100, App
     9   1563.5    38.2    627  9  US-10-238-183-100           Sequence 100, App
    10   1563.5    38.2    627  9  US-10-238-283-100           Sequence 100, App
    11   1563.5    38.2    627  9  US-10-238-370-100           Sequence 100, App
    12   1563.5    38.2    627  9  US-10-245-055-100           Sequence 100, App
    13   1563.5    38.2    627  9  US-10-245-147-100           Sequence 100, App
    14   1563.5    38.2    627  9  US-10-245-730-100           Sequence 100, App
    15   1563.5    38.2    627  9  US-10-245-739-100           Sequence 100, App
    16   1563.5    38.2    627  9  US-10-246-210-100           Sequence 100, App
    17   1563.5    38.2    627  9  US-10-239-196-100           Sequence 100, App
    18   1563.5    38.2    627  9  US-10-243-024-100           Sequence 100, App
    19   1563.5    38.2    627  9  US-10-243-409-100           Sequence 100, App
```

```
    20   1563.5    38.2    627  9  US-10-245-033-100           Sequence 100, App
    21   1563.5    38.2    627  9  US-10-245-621-100           Sequence 100, App
    22   1563.5    38.2    627  9  US-10-245-880-100           Sequence 100, App
    23   1563.5    38.2    627  9  US-10-243-095-100           Sequence 100, App
    24   1563.5    38.2    627  9  US-10-245-185-100           Sequence 100, App
    25   1563.5    38.2    627  9  US-10-245-427-100           Sequence 100, App
    26   1563.5    38.2    627  9  US-10-245-473-100           Sequence 100, App
    27   1563.5    38.2    627  9  US-10-245-770-100           Sequence 100, App
    28   1563.5    38.2    627  9  US-10-245-877-100           Sequence 100, App
    29   1563.5    38.2    627  9  US-10-246-976-100           Sequence 100, App
    30   1563.5    38.2    627  9  US-10-243-320-100           Sequence 100, App
    31   1563.5    38.2    627  9  US-10-242-743-100           Sequence 100, App
    32   1563.5    38.2    627  9  US-10-242-845-100           Sequence 100, App
    33   1545.5    37.8    551 10  US-09-897-214-8               Sequence 8, Appli
    34     1454    35.5    565 10  US-09-768-826-58              Sequence 58, Appl
    35     1144    27.9    468 10  US-09-768-826-40              Sequence 40, Appl
    36      452    11.0    713  9  US-10-028-392-35            Sequence 35, Appl
    37      451    11.0    713 10  US-09-822-687-2               Sequence 2, Appli
    38      441    10.8    634  9  US-10-028-392-2             Sequence 2, Appli
    39    435.5    10.6    640  9  US-09-905-291A-292          Sequence 292, App
    40    435.5    10.6    640  9  US-09-992-598-501           Sequence 501, App
    41    435.5    10.6    640  9  US-09-989-293A-501          Sequence 501, App
    42    435.5    10.6    640  9  US-09-902-853-292           Sequence 292, App
    43    435.5    10.6    640  9  US-09-989-735-501           Sequence 501, App
    44    435.5    10.6    640  9  US-09-990-444-501           Sequence 501, App
    45    435.5    10.6    640  9  US-10-114-893-117           Sequence 117, App

                                ALIGNMENTS


RESULT 1
US-09-815-626-2
; Sequence 2, Application US/09815626
; Patent No. US20020076752A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria Alexandra
;  TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT
;  TITLE OF INVENTION:  FAMILY MEMBER AND USES THEREOF
;  FILE REFERENCE: 10448-031001
;  CURRENT APPLICATION NUMBER: US/09/815,626
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: US 60/191,863
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 628
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-815-626-2

  Query Match             38.3%;  Score 1566;  DB 10;  Length 628;
  Best Local Similarity   52.9%;  Pred. No. 1.2e-78;
  Matches  311;  Conservative   88;  Mismatches  175;  Indels   14;  Gaps    5;

Qy         16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
              |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db         23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy         76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
              |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db         83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy        136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
              |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db        143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASVPAGAFSR 202

Qy        196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
              | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db        203 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 259

Qy        253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
```

```
              |||| |:|||| | ||  |  ||||| | |||||||||::|  :  | |  |: | |:|:|
Db        260 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 319

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db        320 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 379

Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db        380 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 435

Qy        431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                 |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db        436 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 495

Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||::: :|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db        496 VLAVYEGSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 555

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db        556 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 603
```

```
RESULT 2
US-10-245-103-100
; Sequence 100, Application US/10245103
; Publication No. US20030068778A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C112
;  CURRENT APPLICATION NUMBER: US/10/245,103
;  CURRENT FILING DATE:  2002-09-17
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-103-100
```

```
  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy         16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
              |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db         23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy         76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
              |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db         83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy        136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
              |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db        143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy        196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
              | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db        202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy        253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
              |||| |:|||| | ||  |  ||||| | |||||||||::|  :  | |  |: | |:|:|
Db        259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db        319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db        379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy        431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                 |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db        435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||::: :|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db        495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db        555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602
```

```
RESULT 3
US-10-245-107-100
; Sequence 100, Application US/10245107
; Publication No. US20030068779A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C71
;  CURRENT APPLICATION NUMBER: US/10/245,107
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
```

```
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-107-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602


RESULT 4
US-10-245-143-100
; Sequence 100, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C90
;  CURRENT APPLICATION NUMBER: US/10/245,143
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-143-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378
```

```
Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db        379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy        431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                 |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db        435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||:::|:|| ||||  ||||:| |:     | : |:  |||||| : :||:|||::||||
Db        495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : | :|:  | ||:| ||||||  | |: ||    |
Db        555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 5
US-10-245-771-100
; Sequence 100, Application US/10245771
; Publication No. US20030068781A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C98
;  CURRENT APPLICATION NUMBER: US/10/245,771
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-771-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy         16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
              |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db         23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy         76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
              |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db         83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy        136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
              |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db        143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy        196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
              | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db        202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy        253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
              |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db        259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db        319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db        379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy        431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                 |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db        435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||:::|:|| ||||  ||||:| |:     | : |:  |||||| : :||:|||::||||
Db        495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : | :|:  | ||:| ||||||  | |: ||    |
Db        555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 6
US-10-245-851-100
; Sequence 100, Application US/10245851
; Publication No. US20030068782A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C93
;  CURRENT APPLICATION NUMBER: US/10/245,851
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
```

```
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-851-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 7
US-10-245-883-100
; Sequence 100, Application US/10245883
; Publication No. US20030068783A1
; GENERAL INFORMATION:
;  APPLICANT:  Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C70
;  CURRENT APPLICATION NUMBER: US/10/245,883
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-883-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494
```

```
Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db        495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db        555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 8
US-10-237-535-100
; Sequence 100, Application US/10237535
; Publication No. US20030073188A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C3
;  CURRENT APPLICATION NUMBER: US/10/237,535
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/091358
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091978
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/099803
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/106932
;  PRIOR FILING DATE: 1998-11-03
;  PRIOR APPLICATION NUMBER: 60/115554
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/119342
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: 60/123957
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123972
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: 60/127372
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/131271
;  PRIOR FILING DATE: 1999-04-27
;  PRIOR APPLICATION NUMBER: 60/133459
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/135725
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/135729
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/135750
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/138385
;  PRIOR FILING DATE: 1999-06-09
;  PRIOR APPLICATION NUMBER: 60/140653
;  PRIOR FILING DATE: 1999-06-22
;  PRIOR APPLICATION NUMBER: 60/141037
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: 60/144732
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/144758
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/144790
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/145228
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/145698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: 60/146222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: 60/146843
;  PRIOR FILING DATE: 1999-08-03
;  PRIOR APPLICATION NUMBER: 60/148188
;  PRIOR FILING DATE: 1999-08-10
;  PRIOR APPLICATION NUMBER: 60/148513
;  PRIOR FILING DATE: 1999-08-12
;  PRIOR APPLICATION NUMBER: 60/149327
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/149395
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/150114
;  PRIOR FILING DATE: 1999-08-20
;  PRIOR APPLICATION NUMBER: 60/151700
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: 60/151734
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: 60/162506
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: 60/170262
;  PRIOR FILING DATE: 1999-12-09
;  PRIOR APPLICATION NUMBER: 60/177118
;  PRIOR FILING DATE: 2000-01-20
;  PRIOR APPLICATION NUMBER: 60/179851
;  PRIOR FILING DATE: 2000-02-02
;  PRIOR APPLICATION NUMBER: 60/180921
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: 60/187202
;  PRIOR FILING DATE: 2000-03-03
;  PRIOR APPLICATION NUMBER: 60/198587
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: 60/199614
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 60/206330
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/206368
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/209832
;  PRIOR FILING DATE: 2000-06-05
;  PRIOR APPLICATION NUMBER: 60/218371
;  PRIOR FILING DATE: 2000-07-13
;  PRIOR APPLICATION NUMBER: 60/222695
;  PRIOR FILING DATE: 2000-08-02
;  PRIOR APPLICATION NUMBER: 60/229896
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/230621
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/232887
;  PRIOR FILING DATE: 2000-09-15
```

```
;  PRIOR APPLICATION NUMBER: 60/235147
;  PRIOR FILING DATE: 2000-09-22
;  PRIOR APPLICATION NUMBER: 60/261878
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: 60/261910
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/261939
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/262150
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/264395
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/266421
;  PRIOR FILING DATE: 2001-02-02
;  PRIOR APPLICATION NUMBER: 60/267623
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/274399
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 60/280982
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: 60/282129
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: 60/282199
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: 60/290589
;  PRIOR FILING DATE: 2001-05-09
;  PRIOR APPLICATION NUMBER: 09/180997
;  PRIOR FILING DATE: 1998-11-19
;  PRIOR APPLICATION NUMBER: 09/267213
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 09/380137
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/380138
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/403297
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 09/423741
;  PRIOR FILING DATE: 1999-11-10
;  PRIOR APPLICATION NUMBER: 09/709238
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 09/802706
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 09/872035
;  PRIOR FILING DATE: 2001-06-01
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/924419
;  PRIOR FILING DATE: 2001-08-06
;  PRIOR APPLICATION NUMBER: 09/927796
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 09/929404
;  PRIOR FILING DATE: 2001-08-13
;  PRIOR APPLICATION NUMBER: 09/931836
;  PRIOR FILING DATE: 2001-08-16
;  PRIOR APPLICATION NUMBER: 09/941992
;  PRIOR FILING DATE: 2001-08-28
;  PRIOR APPLICATION NUMBER: 09/946374
;  PRIOR FILING DATE: 2001-09-04
;  PRIOR APPLICATION NUMBER: 10/001054
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 10/081056
;  PRIOR FILING DATE: 2002-02-20
;  PRIOR APPLICATION NUMBER: 10/119480
;  PRIOR FILING DATE: 2002-04-09

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy         16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
              |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db         23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy         76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
              |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db         83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy        136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
              |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db        143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy        196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
              | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db        202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy        253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
              |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db        259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||  |
Db        319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db        379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy        431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                 |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db        435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db        495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db        555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 9
US-10-238-183-100
; Sequence 100, Application US/10238183
; Publication No. US20030073189A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C11
;  CURRENT APPLICATION NUMBER: US/10/238,183
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
```

```
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/091358
;  PRIOR FILING DATE: 1998-07-01
;  PRIOR APPLICATION NUMBER: 60/091978
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/099803
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/106932
;  PRIOR FILING DATE: 1998-11-03
;  PRIOR APPLICATION NUMBER: 60/115554
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/119342
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: 60/123957
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123972
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: 60/127372
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/131271
;  PRIOR FILING DATE: 1999-04-27
;  PRIOR APPLICATION NUMBER: 60/133459
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/135725
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/135729
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/135750
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 60/138385
;  PRIOR FILING DATE: 1999-06-09
;  PRIOR APPLICATION NUMBER: 60/140653
;  PRIOR FILING DATE: 1999-06-22
;  PRIOR APPLICATION NUMBER: 60/141037
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: 60/144732
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/144758
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/144790
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/145228
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/145698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: 60/146222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: 60/146843
;  PRIOR FILING DATE: 1999-08-03
;  PRIOR APPLICATION NUMBER: 60/148188
;  PRIOR FILING DATE: 1999-08-10
;  PRIOR APPLICATION NUMBER: 60/148513
;  PRIOR FILING DATE: 1999-08-12
;  PRIOR APPLICATION NUMBER: 60/149327
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/149395
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/150114
;  PRIOR FILING DATE: 1999-08-20
;  PRIOR APPLICATION NUMBER: 60/151700
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: 60/151734
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: 60/162506
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: 60/170262
;  PRIOR FILING DATE: 1999-12-09
;  PRIOR APPLICATION NUMBER: 60/177118
;  PRIOR FILING DATE: 2000-01-20
;  PRIOR APPLICATION NUMBER: 60/179851
;  PRIOR FILING DATE: 2000-02-02
;  PRIOR APPLICATION NUMBER: 60/180921
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: 60/187202
;  PRIOR FILING DATE: 2000-03-03
;  PRIOR APPLICATION NUMBER: 60/198587
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: 60/199614
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 60/206330
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/206368
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/209832
;  PRIOR FILING DATE: 2000-06-05
;  PRIOR APPLICATION NUMBER: 60/218371
;  PRIOR FILING DATE: 2000-07-13
;  PRIOR APPLICATION NUMBER: 60/222695
;  PRIOR FILING DATE: 2000-08-02
;  PRIOR APPLICATION NUMBER: 60/229896
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/230621
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/232887
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: 60/235147
;  PRIOR FILING DATE: 2000-09-22
;  PRIOR APPLICATION NUMBER: 60/261878
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: 60/261910
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/261939
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/262150
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/264395
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/266421
;  PRIOR FILING DATE: 2001-02-02
;  PRIOR APPLICATION NUMBER: 60/267623
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/274399
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 60/280982
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: 60/282129
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: 60/282199
;  PRIOR FILING DATE: 2001-04-04
;  PRIOR APPLICATION NUMBER: 60/290589
;  PRIOR FILING DATE: 2001-05-09
;  PRIOR APPLICATION NUMBER: 09/180997
;  PRIOR FILING DATE: 1998-11-19
;  PRIOR APPLICATION NUMBER: 09/267213
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 09/380137
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/380138
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/403297
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 09/423741
;  PRIOR FILING DATE: 1999-11-10
;  PRIOR APPLICATION NUMBER: 09/709238
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 09/802706
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 09/872035
;  PRIOR FILING DATE: 2001-06-01
```

```
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/924419
;  PRIOR FILING DATE: 2001-08-06
;  PRIOR APPLICATION NUMBER: 09/927796
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 09/929404
;  PRIOR FILING DATE: 2001-08-13
;  PRIOR APPLICATION NUMBER: 09/931836
;  PRIOR FILING DATE: 2001-08-16
;  PRIOR APPLICATION NUMBER: 09/941992
;  PRIOR FILING DATE: 2001-08-28
;  PRIOR APPLICATION NUMBER: 09/946374
;  PRIOR FILING DATE: 2001-09-04
;  PRIOR APPLICATION NUMBER: 10/001054
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 10/081056
;  PRIOR FILING DATE: 2002-02-20
;  PRIOR APPLICATION NUMBER: 10/119480
;  PRIOR FILING DATE: 2002-04-09

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | ||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:|||  ||::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 10
US-10-238-283-100
; Sequence 100, Application US/10238283
; Publication No. US20030073190A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C15
;  CURRENT APPLICATION NUMBER: US/10/238,283
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-238-283-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | ||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378
```

```
Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 11
US-10-238-370-100
; Sequence 100, Application US/10238370
; Publication No. US20030073191A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION:  SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C10
;  CURRENT APPLICATION NUMBER: US/10/238,370
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-238-370-100

  Query Match            38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity  53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 12
US-10-245-055-100
; Sequence 100, Application US/10245055
; Publication No. US20030073192A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION:  SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C88
;  CURRENT APPLICATION NUMBER: US/10/245,055
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
```

```
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-055-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

RESULT 13
US-10-245-147-100
; Sequence 100, Application US/10245147
; Publication No. US20030073193A1
; GENERAL INFORMATION:
;  APPLICANT:  Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C72
;  CURRENT APPLICATION NUMBER: US/10/245,147
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File.Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-147-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494
```

```
Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db        495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db        555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602


RESULT 14
US-10-245-730-100
; Sequence 100, Application US/10245730
; Publication No. US20030073194A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C85
;  CURRENT APPLICATION NUMBER: US/10/245,730
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-730-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy         16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
              |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db         23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy         76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
              |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db         83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy        136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
              |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db        143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy        196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
              | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db        202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy        253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
              |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db        259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db        319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy        373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db        379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy        431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                 |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db        435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy        491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
              |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db        495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy        551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
               :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db        555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602


RESULT 15
US-10-245-739-100
; Sequence 100, Application US/10245739
; Publication No. US20030073195A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin
;  APPLICANT:  Eaton,Dan
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Grimaldi,J. Christopher
;  APPLICANT:  Gurney,Austin
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stephan,Jean-Phillippe
;  APPLICANT:  Watanbe,Colin
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang,Zemin
;  APPLICANT:  Fong,Sherman
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3630R1C86
;  CURRENT APPLICATION NUMBER: US/10/245,739
;  CURRENT FILING DATE:  2002-09-16
;  PRIOR APPLICATION NUMBER: 10/197942
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: 60/059114
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/063046
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/065027
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/086478
;  PRIOR FILING DATE: 1998-05-22
;  PRIOR APPLICATION NUMBER: 60/087607
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/089801
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/090557
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090689
```

```
;  PRIOR FILING DATE: 1998-06-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-245-739-100

  Query Match             38.2%;  Score 1563.5;  DB 9;  Length 627;
  Best Local Similarity   53.1%;  Pred. No. 1.7e-78;
  Matches  312;  Conservative   88;  Mismatches  173;  Indels   15;  Gaps    6;

Qy          16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
               |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db          23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy          76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
               |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db          83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy         136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
               |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db         143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASV-PGAFSR 201

Qy         196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
               | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db         202 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 258

Qy         253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
               |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db         259 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 318

Qy         313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
               :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db         319 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378

Qy         373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
               |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db         379 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434

Qy         431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                  |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db         435 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 494

Qy         491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
               |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db         495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 554

Qy         551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db         555 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 602

Search completed: June  8, 2003, 21:38:55
Job time : 59 secs
```

```
                        GenCore version 5.1.6
               Copyright (c) 1993 - 2003  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         June  8, 2003, 21:20:53 ; Search time 44 Seconds
                                          (without alignments)
                                          1723.865 Million cell updates/sec

Title:          US-09-831-846-2
Perfect score:  4094
Sequence:       1 METLLGGLLAFGMAFAVVDA..........DLVGARGTFGSSEWVMESTV 789

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters:      283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_73:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                      %
Result                Query
  No.    Score    Match Length DB  ID                        Description
-------------------------------------------------------------------------------
    1    380.5      9.3    707  2  JC7763                    neuronal leucine-r
    2      368      9.0   1091  2  A58532                    glial cell membran
    3      350      8.5    421  2  T46266                    hypothetical prote
    4    286.5      7.0   1531  2  T42218                    slit-1 protein hom
    5      263      6.4   1469  2  B36665                    slit protein 2 pre
    6      263      6.4   1480  2  A36665                    slit protein 1 pre
    7      259      6.3   1523  2  T13953                    MEGF5 protein - ra
    8      254      6.2    789  2  T28714                    hypothetical prote
    9      254      6.2   1355  2  T28715                    hypothetical prote
   10      248      6.1    603  2  JC1282                    insulin-like growt
   11    241.5      5.9    605  2  A41915                    insulin-like growt
   12      237      5.8   1025  2  T42626                    secreted leucine-r
   13      236      5.8    361  2  A53860                    chondroadherin pre
   14    235.5      5.8    605  2  JC5239                    insulin-like growt
   15    232.5      5.7   1535  2  S46224                    peroxidasin - frui
   16      231      5.6    603  2  JC6128                    insulin-like growt
   17    226.5      5.5   1328  2  T23007                    hypothetical prote
   18      221      5.4    907  2  JE0176                    orphan G protein-c
   19      208      5.1    907  2  JG0193                    G protein-coupled
   20      208      5.1   1896  2  T08851                    Down syndrome cell
   21    207.5      5.1   1066  2  T15864                    hypothetical prote
   22    203.5      5.0    738  2  T19938                    hypothetical prote
   23      203      5.0    821  1  S06943                    brain-derived neur
   24    202.5      4.9   1389  2  T13852                    gene wheeler prote
   25      202      4.9    333  2  T34555                    hypothetical prote
   26    201.5      4.9    420  2  A53531                    oncofetal trophobl
   27      201      4.9    626  1  NBHUIA                    platelet glycoprot
   28      201      4.9   1385  2  T13887                    tlr protein - frui
   29      198      4.8    536  2  A34901                    lysine carboxypept
```

```
   30      197      4.8    560  2  A60164                    platelet membrane
   31      196      4.8    821  1  A39667                    brain-derived neur
   32    195.5      4.8    458  2  T19941                    hypothetical prote
   33    195.5      4.8   1344  2  T14316                    rig-1 protein - mo
   34      195      4.8    360  2  S06280                    decorin precursor
   35      195      4.8   1097  2  A29943                    Toll protein precu
   36      193      4.7    562  2  T34319                    hypothetical prote
   37      193      4.7    653  2  T25194                    hypothetical prote
   38      191      4.7    702  2  T21148                    hypothetical prote
   39    190.5      4.7    961  2  T23395                    hypothetical prote
   40    189.5      4.6    476  1  A35104                    brain-derived neur
   41    188.5      4.6   1257  2  A88536                    protein B0523.5 [i
   42    186.5      4.6    375  2  S05390                    fibromodulin precu
   43    186.5      4.6   1134  1  A29944                    chaoptin precursor
   44    185.5      4.5    354  2  A55454                    decorin precursor
   45    185.5      4.5    369  2  S20811                    proteoglycan I - m


                              ALIGNMENTS


RESULT 1
JC7763
neuronal leucine-rich repeat protein-3 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: JC7763
R;Fukamachi, K.; Matsuoka, Y.; Kitanaka, C.; Kuchino, Y.; Tsuda, H.
Biochem. Biophys. Res. Commun. 287, 257-263, 2001
A;Title: Rat neuronal leucine-rich repeat protein-3: Cloning and regulation of the gene
A;Reference number: JC7763; PMID:11549284
A;Contents: Fibrosarcoma cells
A;Accession: JC7763
A;Molecule type: mRNA
A;Residues: 1-707 <FUK>
A;Cross-references: GB:AF291437
C;Comment: This protein, a new member of the neuronal leucine-rich repeat protein family
in protein-protein interaction and functions as a cell adhesion molecule or soluble liga
C;Genetics:
A;Gene: nlrr-3
C;Keywords: cell adhesion

  Query Match             9.3%;  Score 380.5;  DB 2;  Length 707;
  Best Local Similarity   22.4%;  Pred. No. 3.5e-17;
  Matches  154;  Conservative   97;  Mismatches  233;  Indels  203;  Gaps   27;

Qy         12 GMAF-AVVDA------CPKYCVCQ--------NLSESLGTL-CPSKGLLFVPPDIDRRTV 55
              |:|  |:| |      ||: | |:        ::     |: |   ||| |  :   |
Db         13 GLAITALVQAGDKKVDCPQLCTCEIRPWFTPRSIYMEASTVDCNDLGLLNFPARLPADTQ 72

Qy         56 ELRLGGNFIIHISRQ-DF-ANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLP 113
               | |  | |  |    ||  |:||   | |:| :| :   :   :  | |::|: |:|
Db         73 ILLLQTNNIARIEHSTDFPVNLTG---LDLSQNNLSSVTNINVQKMSQLLSVYLEENKLT 129

Qy        114 SLGEDTLRGLVNLQHLIVNNNQLGGIADEAF----------------------------- 144
               | |  | || ||| | ||:| |  |:  ||
Db        130 ELPEKCLYGLSNLQELYVNHNLLSAISPGAFVGLHNLLRLHLNSNRLQMINSKWFEALPN 189

Qy        145 ---------------------------------------EDFLLTLEDLD-LSY--NNLH 162
                                                     :| |: ||:|: :|:  | |:
Db        190 LEILMLGDNPILRIKDMNFQPLLKLRSLVIAGINLTEVPDDALVGLENLESISFYDNRLN 249

Qy        163 GLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFA-------------------------DLQ 197
               :|  :::: |||  | |: | :: |  |  |:                         :|
Db        250 KVPQVALQKAVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELVSIDSLAVDNLP 309

Qy        198 KLARLDLTSN-RLQKLPPDPIFARSQASALTATPFA------------PPL-SFSFGGNP 243
               | ::: |:| ||  : |:  |   :  :|     |            | |   |   ||
Db        310 DLRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSNP 369

Qy        244 LHCNCELLWLR------RLERDDDLETCGSPGGLKGRYFWHVREEEF-----VCEPPLIT 292
              : |:| : |:       |    | |  |  |   :|:   :||:  |     :| | :
Db        370 IRCDCVIRWINMNKTNIRFMEPDSL-FCVDPPEFQGQ---NVRQVHFRDMMEICLPLIAP 425
```

```
Qy         293 QHTHKLLVLEGQA-ATLKCKAIGDPSPLIHWVAPD-DRLVGNSSRTA--VYDNGTLDIFI 348
               :    :| :|  :  :| |:|  :| | |:|: |   ||: |: |    |:  |||||
Db         426 ESFPSILDVEADSYVSLHCRATAEPQPEIYWITPSGKRLLPNTLREKFYVHSEGTLDIRG 485

Qy         349 TTSQDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRG 408
                | :: | :|||| |  |     : : :
Db         486 ITPKEGGLYTCIATNLVGADLKSIMIKV-------------------------------- 513

Qy         409 GGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSV-SKSAPRVKMYQLQYNCSDDEVLIYR 467
                   ||  |  :     : : ::   | || |   ||       :       | :
Db         514 ----GGFVPQDNNGSLNIKIRDIRANSVLVSWKANSKILKSSVKWTAFVKTEDSQAAQSA 569

Qy         468 MIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKA---DYPQCQS 524
                ||:  | : : :|   | | :|:         |:   :   |    ||:   |  :
Db         570 RIPSDVKVYNLTHLKPSTEYKICI------DIPTIYQKSRKQCVNVTTKSLEHDGKENGK 623

Qy         525 MHSQILGGTMILVIGGI--IVATLLVF 549
                |:       :  :||:  |:  : :|
Db         624 SHT-----VFVACVGGLLGIIGVMCLF 645

RESULT 2
A58532
glial cell membrane glycoprotein LIG-1 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 05-Nov-1999
C;Accession: A58532
R;Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.
J. Biol. Chem. 271, 22522-22527, 1996
A;Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically i
A;Reference number: A58532; MUID:96394313; PMID:8798419
A;Accession: A58532
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1091 <SUZ>
A;Cross-references: GB:D78572; NID:g1545806; PIDN:BAA11416.1; PID:g1545807
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan amino-ter
F;36-61/Domain: proteoglycan amino-terminal homology <PAH>
F;71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
F;310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
F;334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F;358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
F;385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F;409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F;440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

  Query Match             9.0%;  Score 368;  DB 2;  Length 1091;
  Best Local Similarity   21.3%;  Pred. No. 4.2e-16;
  Matches  205;  Conservative   92;  Mismatches  326;  Indels  340;  Gaps   33;

Qy          30 LSESLGTLCPSKGLL-------FVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVDLT 82
               || || ||  ||  :     . |  |    |  :| |  | |  |    |  :  |  |
Db         188 LSRSLLTLRLSKNRITQLPVKAFKLP----RLTQLDLNRNRIRLIEGLTFQGLDSLEVLR 243

Qy          83 LSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADE 142
               | || || :   :|  |  :  |||: | |  :   :| ||  |  | ::|| :  |  :
Db         244 LQRNNISRLTDGAFWGLSKMHVLHLEYNSLVEVNSGSLYGLTALHQLHLSNNSISRIQRD 303

Qy         143 AFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLARL 202
                :  |   | :| ||:|||  |  :|:  : :|  | | || : ||||| |  |: |  |
Db         304 GW-SFCQKLHELILSFNNLTRLDEESLAELSSLSILRLSHNAISHIAEGAFKGLKSLRVL 362

Qy         203 DLTSNRLQKLPPDPIFARSQASALT-ATPFAPPLS---------------FSFGGNPLH- 245
               ||  | :     |   | :    |:  | |   :                 : | | :
Db         363 DLDHNEISGTIEDTSGAFTGLDNLSKLTLFGNKIKSVAKRAFSGLESLEHLNLGENAIRS 422

Qy         246 -----------------------CNCELLWL-----RRLERDDDLETCGSPGGLKGRYFW 277
                                      |:|:| ||      |: :     ||  |  |||:  :
Db         423 VQFDAFAKMKNLKELYISSESFLCDCQLKWLPPWLMGRMLQAFVTATCAHPESLKGQSIF 482

Qy         278 HVREEEFVCE----PPLITQ---------------------------------------- 293
                |  : |||:    | :|||
Db         483 SVLPDSFVCDDFPKPQIITQPETTMAVVGKDIRFTCSAASSSSPMTFAWKKDNEVLANA 542

Qy         294 ----------------------------------------------HTHK---------- 297
                                                             ::||
Db         543 DMENFAHVRAQDGEVMEYTTILHLRHVTFGHEGRYQCIITNHFGSTYSHKARLTVNVLPS 602

Qy         298 -------LLVLEGQAATLKCKAIGDPSPLIHW-----------------VAPDDRLVGNS 333
                      : :  |  | |:| | | |:| | |                 | |||
Db         603 FTKIPHDIAIRTGTTARLECAATGHPNPQIAWQKDGGTDFPAARERRMHVMPDD------ 656

Qy         334 SRTAVYDNGTLDIFITTS---QDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTA 390
                          |:|  |     | | ::| | |:||  :|   ::::: | |       |
Db         657 -----------DVFFITDVKIDDMGVYSCTAQNSAGSVSANATLTVLETPSL-------A 698

Qy         391 PPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVK 450
                |                              :| | | |   | |    :     ||:
Db         699 VPLE----------------------------DRVVTVGE---TVAFQCKATGSPTPRIT 727

Qy         451 MYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTAT-TLTATNIVG 509
                 :     |  |   :   |  |:  || |:            | ||    |   :| :|
Db         728 WLKGGRPLSLTE--RHHFTP-GNQLLVVQNV------------MIDDAGRYTCEMSNPLG 772

Qy         510 CAQFFTKADY---PQCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRYKVCNHEAPS 566
                 :  ::      | |:      :|   | |:  |:: :|:   :|   | |
Db         773 TERAHSQLSILPTPGCRK-DGTTVGIFTIAVVCSIVLTSLVWVCIIYQTRKK-------- 823

Qy         567 KMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGP----PKVVVRNE--------------- 607
                     |  || ||  :   |   |:    ||      : ||| |
Db         824 ------SEEYSVTNTDETIVPPDVPSYLSSQGTLSDRQETVVRTEGGHQANGHIESNGVC 877

Qy         608 LLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPR-PKPSLDRLMGAF----AS 662
               | |   ||    |  |::        |  | ||::   | |   |      |:      |
Db         878 LRD--PSLFPEVDIHSTTCRQPKLCVGYTREPWKVTEKADRTAAPHTTAHSGSAVCSDCS 935

Qy         663 LDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKRSH 722
                |     :    ||  | |  :    | |  :  |               |  | |  ||
Db         936 TDTAYHPQPVPRDSGQP-GTASSQELRQHDREYSP-------------HHPYSGTADGSH 981

Qy         723 SFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSE 782
               :         : | : |  :                    :||  ::    | |  | |
Db         982 TL--------SGGSLYPSNHD------------------RILPSLKNKAASADGN-GDSS 1014

Qy         783 WVM 785
               | :
Db        1015 WTL 1017

RESULT 3
T46266
hypothetical protein DKFZp761A179.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46266
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23034
A;Accession: T46266
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-421 <AAA>
A;Cross-references: EMBL:AL137451
A;Experimental source: adult amygdala; clone DKFZp761A179
C;Genetics:
A;Note: DKFZp761A179.1
```

```
  Query Match             8.5%;  Score 350;  DB 2;  Length 421;
  Best Local Similarity   27.7%;  Pred. No. 1.7e-15;
  Matches   99;  Conservative   43;  Mismatches  101;  Indels  114;  Gaps    9;

Qy          70 QDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHL 129
               :|  |:| ||                      |  | |  |||  :   : :|| :|: |
Db           5 KDIPNLTALV---------------------RLEELELSGNRLDLIRPGSFQGLTSLRKL 43

Qy         130 IVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIA 189
                : : |:  |   ||:| | :||:|:||:|||  || |
Db          44 WLMHAQVATIERNAFDD-LKSLEELNLSHNNLMSLPHD---------------------- 80

Qy         190 EGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCE 249
                  |  | :| |: |                                     || ||||:
Db          81 --LFTPLHRLERVHLNH-----------------------------------NPWHCNCD 103

Qy         250 LLWLR-RLERDDDLET-----CGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEG 303
               :|||   |:      |     | :| ||||||   : :  | |  |:| :    | | ||
Db         104 VLWLSWWLKETVPSNTTCCARCHAPAGLKGRYIGELDQSHFTCYAPVIVEPPTDLNVTEG 163

Qy         304 QAATLKCKAIGDPSPLIHWVAPDDRLVGNSS---RTAVYDNGTLDIFITTSQDSGAFTCI 360
                || |||:  |     ::|: |: |: : |   | :|  :|||:    | ||:| :||:
Db         164 MAAELKCRT-GTSMTSVNWLTPNGTLMTHGSYRVRISVLHDGTLNFTNVTVQDTGQYTCM 222

Qy         361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEP 417
                 |:||  ||                       |   :::     : ||  ||||| |
Db         223 VTNSAGNTTA-----------------------SATLNVSAVDPVAAGGTGSGGGGP 256


RESULT 4
T42218
slit-1 protein homolog - rat
N;Alternate names: MEGF4 protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2002
C;Accession: T42218
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A;Reference number: Z14126; MUID:98360089; PMID:9693030
A;Accession: T42218
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1531 <NAK>
A;Cross-references: EMBL:AB011530; NID:g3449289; PIDN:BAA32460.1; PID:g3449290
A;Experimental source: strain Sprague-Dawley; brain
C;Genetics:
A;Gene: MEGF4
C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r

  Query Match             7.0%;  Score 286.5;  DB 2;  Length 1531;
  Best Local Similarity   30.3%;  Pred. No. 1.5e-10;
  Matches   87;  Conservative   33;  Mismatches  116;  Indels   51;  Gaps    8;

Qy           8 LLAFGMAFAV-VDACPKYCVCQNLSESLGTL--CPSKGLLFVPPDIDRRTVELRLGGNFI 64
               || :  |: :   |||  | |       ||   |   ||  :| :| | |  | | || |
Db          20 LLLWAAAWRLGATACPALCTC------TGTTVDCHGTGLQAIPKNIPRNTERLELNGNNI 73

Qy          65 IHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLV 124
                 | : ||| :  |  | |  | |  ::  :| |:: |  | |: |:|  | |
Db          74 TRIHKNDFAGLKQLRVLQLMENQIGAVERGAFDDMKELERLRLNRNQLQVLPE------- 126

Qy         125 NLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNL 184
                   |:  |||               |  |||| |:|  :|   :| |   :|  | || |
Db         127 ----LLFQNNQ--------------ALSRLDLSENSLQAVPRKAFRGATDLKNLQLDKNQ 168

Qy         185 LDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPL-SFSFGGNP 243
               :  | || | |:| |  | | :| : :|            :::    | :|    |
Db         169 ISCIEEGAFRALRGLEVLTLNNNNITTIP------------VSSFNHMPKLRTFRLHSNH 216

Qy         244 LHCNCELLWLRRLERDDD----LETCGSPGGLKGRYFWHVREEEFVC 286
               | |:| | || :  |          |  |  |:|     |:: || |
```

```
Db         217 LFCDCHLAWLSQWLRQRPTIGLFTQCSGPASLRGLNVAEVQKSEFSC 263

RESULT 5
B36665
slit protein 2 precursor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
C;Accession: B36665
R;Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2169-2187, 1990
A;Title: slit: an extracellular protein necessary for development of midline glia and co
A;Reference number: A36665; MUID:91099665; PMID:2176636
A;Accession: B36665
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1469 <ROT>
A;Cross-references: GB:X53959
C;Genetics:
A;Gene: FlyBase:sli
A;Cross-references: FlyBase:FBgn0003425
C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r
F;66-91/Domain: proteoglycan amino-terminal homology <PAH1>
F;101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>
F;288-313/Domain: proteoglycan amino-terminal homology <PAH2>
F;323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
F;450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>
F;512-537/Domain: proteoglycan amino-terminal homology <PAH3>
F;547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
F;572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F;596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
F;620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F;651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
F;708-733/Domain: proteoglycan amino-terminal homology <PAH4>
F;743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F;767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
F;846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
F;1028-1061/Domain: EGF homology <EGF>
F;1068-1099/Domain: EGF homology <EGF2>
F;1115-1148/Domain: EGF homology <EGF1>

  Query Match             6.4%;  Score 263;  DB 2;  Length 1469;
  Best Local Similarity   28.0%;  Pred. No. 5.1e-09;
  Matches   89;  Conservative   39;  Mismatches  102;  Indels   88;  Gaps   11;

Qy           6 GGLLAFGMAF------AVVDA-CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELR 58
               ||| : |:         : :| ||: | |  |:      |  :||  ||  |      |
Db          51 GGLGSVGIHIPGGGVGVITEARCPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106

Qy          59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
               | || :  |   ||  :| |  | |: | |  |:  || || |  | : :| : ::|
Db         107 LQGNNLTVIYETDFQRLTKLRMLQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRR 166

Qy         119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                 :|  :|: | ::|||:  : : ||:  |: || | |: |||  ||
Db         167 VFKGAQSLRSLQLDNNQITCLDEHAFKG-LVELEILTLNNNNLTSLP------------- 212

Qy         179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
                  ||:        |  | :|  | |: |                      |||
Db         213 ---HNI--------FGGLGRLRALRLSDN----------------------PFA------ 233

Qy         239 FGGNPLHCNCELLWLRRLERD----DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                      |:| | || |  |          | ||  |||:    : :||| |    :|:|
Db         234 -------CDCHLSWLSRFLRSATRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 284
```

```
Qy         295 THKLLVLEGQAATLKCKA 312
                          | ::| |
Db         285 -----------APMECGA 291

RESULT 6
A36665
slit protein 1 precursor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
C;Accession: A36665; A31640; S13523
R;Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2169-2187, 1990
A;Title: slit: an extracellular protein necessary for development of midline glia and c
A;Reference number: A36665; MUID:91099665; PMID:2176636
A;Accession: A36665
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1480 <ROT>
A;Cross-references: GB:X53959; NID:g8614; PIDN:CAA37910.1; PID:g8615
R;Rothberg, J.M.; Hartley, D.A.; Walther, Z.; Artavanis-Tsakonas, S.
Cell 55, 1047-1059, 1988
A;Title: slit: An EGF-homologous locus of D. melanogaster involved in the development of
A;Reference number: A31640; MUID:89077533; PMID:3144436
A;Accession: A31640
A;Molecule type: DNA
A;Residues: 881-1182,'G',1185-1404,'GT',1463-1464,'YHA' <RO2>
A;Cross-references: GB:M23543; NID:g340939; PID:g514357
C;Genetics:
A;Gene: FlyBase:sli
A;Cross-references: FlyBase:FBgn0003425
A;Introns: 1351/3
C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r
C;Keywords: alternative splicing; growth factor
F;66-91/Domain: proteoglycan amino-terminal homology <PAH1>
F;101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>
F;288-313/Domain: proteoglycan amino-terminal homology <PAH2>
F;323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
F;450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>
F;512-537/Domain: proteoglycan amino-terminal homology <PAH3>
F;547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
F;572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F;596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
F;620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F;651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
F;708-733/Domain: proteoglycan amino-terminal homology <PAH4>
F;743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F;767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
F;791-814/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
F;815-838/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
F;846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
F;1028-1061/Domain: EGF homology <EGF>
F;1068-1099/Domain: EGF homology <EGF2>
F;1115-1148/Domain: EGF homology <EGF1>

  Query Match             6.4%;  Score 263;  DB 2;  Length 1480;
  Best Local Similarity   28.0%;  Pred. No. 5.2e-09;
  Matches   89;  Conservative   39;  Mismatches  102;  Indels   88;  Gaps   11;

Qy           6 GGLLAFGMAF------AVVDA-CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELR 58
               ||| : |:         : :| ||: | |  |:      |  :||  ||  |      |
Db          51 GGLGSVGIHIPGGGVGVITEARCPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106
```

```
Qy          59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
               | || :  |   ||  :| |  | |: | |  |:  || || ||  | : :| : ::|
Db         107 LQGNNLTVIYETDFQRLTKLRMLQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRR 166

Qy         119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                 :|  :|: | ::|||:  : : ||:  |: || | |: |||  ||
Db         167 VFKGAQSLRSLQLDNNQITCLDEHAFKG-LVELEILTLNNNNLTSLP------------- 212

Qy         179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
                  ||:        |  | :|  | |: |                      |||
Db         213 ---HNI--------FGGLGRLRALRLSDN----------------------PFA------ 233

Qy         239 FGGNPLHCNCELLWLRRLERD----DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                      |:| | || |  |          | || .|||:    : ::|| |    :|:|
Db         234 -------CDCHLSWLSRFLRSATRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 284

Qy         295 THKLLVLEGQAATLKCKA 312
                          | ::| |
Db         285 -----------APMECGA 291

RESULT 7
T13953
MEGF5 protein - rat
N;Alternate names: slit protein homolog
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
C;Accession: T13953
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A;Reference number: Z14126; MUID:98360089; PMID:9693030
A;Accession: T13953
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1523 <NAK>
A;Cross-references: EMBL:AB011531; NID:g3449291; PIDN:BAA32461.1; PID:g3449292
C;Genetics:
A;Gene: MEGF5
C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r

  Query Match             6.3%;  Score 259;  DB 2;  Length 1523;
  Best Local Similarity   27.3%;  Pred. No. 9.9e-09;
  Matches   74;  Conservative   35;  Mismatches   94;  Indels   68;  Gaps    4;

Qy          20 ACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLV 79
               :||  | | |        |  |||  :| ::    ||:||  | |  |    |     |
Db         279 SCPSACSCSNNIVD----CRGKGLTEIPANLPEGIVEIRLEQNSIKSIPAGAFIQYKKLK 334

Qy          80 DLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGI 139
                : :|:| || | | :|  |:|| || |  |::  : :    |||:|| |::| |::  :
Db         335 RIDISKNQISDIAPDAFQGLKSLTSLVLYGNKITEIPKGLFDGLVSLQLLLLNANKINCL 394

Qy         140 ADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKL 199
                   |:|                         : ||: |||  | |  |::| || || :
Db         395 RVNTFQD-------------------------LQNLNLLSLYDNKLQTISKGLFAPLQSI 429

Qy         200 ARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLERD 259
                 | |                                     ||  |:| | ||    :|
Db         430 QTLHLAQ-----------------------------------NPFVCDCHLKWLADYLQD 454

Qy         260 DDLET----CGSPGGLKGRYFWHVREEEFVC 286
               : :||    | ||  |  :    :: ::| |
Db         455 NPIETSGARCSSPRRLANKRISQIKSKKFRC 485

RESULT 8
T28714
hypothetical protein T21D12.9a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28714
R;Woessner, J.
```

```
submitted to the EMBL Data Library, August 1997
A;Description: The sequence of C. elegans cosmid T21D12.
A;Reference number: Z20514
A;Accession: T28714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-789 <WOE>
A;Cross-references: EMBL:AF016687; PIDN:AAC48096.1; GSPDB:GN00022; CESP:T21D12.9a
A;Experimental source: strain Bristol N2; clone T21D12
C;Genetics:
A;Gene: CESP:T21D12.9a
A;Map position: 4
A;Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2

  Query Match             6.2%;  Score 254;  DB 2;  Length 789;
  Best Local Similarity   21.3%;  Pred. No. 8.5e-09;
  Matches  131;  Conservative   75;  Mismatches  204;  Indels  204;  Gaps   25;

Qy          56 ELRLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDS------ 109
               :| |  | |  |    |::   || ||:|| |: :  |||  |  | || |
Db          83 KLDLASNSITDIGTDHFSSFNTLVTLKLARNHITTLNQFSFSRLRKLESLDLTRNMIREV 142

Qy         110 -----NRLPSLGEDTLR-------------GLVNLQHLIVNNNQLGGIADEAFEDFLLTL 151
                    |:||||   :|                  |:|| :: |::  :  | :   | :|
Db         143 RFLAFNQLPSLQNVSLARNDVYRLDDGMFYACEGLKHLNLSTNRVQAVT-EGWMFGLTSL 201

Qy         152 EDLDLSYN-------------------NLH-----GLPWDSVRRMVNLHQLSLDHNLLDH 187
               | ||||||                   :||      ||  | | :  | :| |  | :|
Db         202 EVLDLSYNQIQSFHISSWSHTPKLKWLSLHSNRIQSLPSGSFRVLRQLEELILSANSIDS 261

Qy         188 IAEGTFADLQKLARLDLTSNRLQ---------------------------KLPPDPIFAR 220
               : :     :  | :|||:|| |                            :: |   | |
Db         262 LHKFALVGMSSLHKLDLSSNTLAVCVEDGAVLYNTSMPFLRSLRFTNNQLRVIPKRAFER 321

Qy         221 SQASA---LTATPFA-------PPLSFS---FGGNPLHCNCELLWLR------RLERDDD 261
                 |     ||  | |        ||         : : |:|:: ||       :|::
Db         322 FPALEELDLTDNPIATIHPEAFEPLELKRLVMNSSSILCDCQISWLASWIYRLKLDKSSI 381

Qy         262 LETCGSPGGLKGRYFWHVREEEFVC--EPP--LITQHTHKLLVLEGQAATLKCKAIGDPS 317
               :  |  |  |   |   :      |  : |   | :   ::  | |: |   |   |  |
Db         382 IAKCSYPPPLADLYVVAIDTANLTCHNDSPRAKIVRQPVEVSTLIGEKARFTCNVYG-AS 440

Qy         318 PL-IHW----------VAPDDRLVGNSSRTAVYDNGTLD----------IFITTSQDSGA 356
               || | |          :  |     : :||||  ||| |          :      |:
Db         441 PLSIEWRVMENGQPRVLVQDSATFLSINRTAVV-NGTFDERELAAAELLLDNVAMTDNSE 499

Qy         357 FTCIAANAAG-EATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGG 415
               : |:| |  | : :  |:: : | |  :
Db         500 YQCVARNRFGSDFSTHVKLQVYQAPKFT-------------------------------- 527

Qy         416 EPPKSPPERAVLVSEVT--------TTSALVKWSVSK-SAPRVKMYQLQYNCSDDEVLIY 466
                   :| :  :|| :          |    :||:  :   |  :  :|    :|| :
Db         528 ---YTPEDMPLLVGQTAKFLCAATGTPRPEIKWAFEQIPFPAAEARRLYVTPNDDHI--- 581

Qy         467 RMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQ----------FFTK 516
                        :::|            :   |  | |  |||: |  |          ||
Db         582 ---------YIMN------------VTKEDQGAYTCHATNVAGQTQASANLIVFENFF-- 618

Qy         517 ADYPQCQSMHSQIL 530
                 ||:   :   ::
Db         619 -HYPESPDLSPMLI 631

RESULT 9
T28715
hypothetical protein T21D12.9b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28715
R;Woessner, J.
submitted to the EMBL Data Library, August 1997
A;Description: The sequence of C. elegans cosmid T21D12.
A;Reference number: Z20514
```

```
A;Accession: T28715
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1355 <WOE>
A;Cross-references: EMBL:AF016687; PIDN:AAC48095.1; GSPDB:GN00022; CESP:T21D12.9b
A;Experimental source: strain Bristol N2; clone T21D12
C;Genetics:
A;Gene: CESP:T21D12.9b
A;Map position: 4
A;Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2; 786/3; 84

  Query Match             6.2%;  Score 254;  DB 2;  Length 1355;
  Best Local Similarity   21.3%;  Pred. No. 1.8e-08;
  Matches  131;  Conservative   75;  Mismatches  204;  Indels  204;  Gaps   25;

Qy          56 ELRLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDS------ 109
               :| |  | |  |    |::   || ||:|| |: :  |||  |  | || |
Db          83 KLDLASNSITDIGTDHFSSFNTLVTLKLARNHITTLNQFSFSRLRKLESLDLTRNMIREV 142

Qy         110 -----NRLPSLGEDTLR-------------GLVNLQHLIVNNNQLGGIADEAFEDFLLTL 151
                    |:||||   :|                  |:|| :: |::  :  | :   | :|
Db         143 RFLAFNQLPSLQNVSLARNDVYRLDDGMFYACEGLKHLNLSTNRVQAVT-EGWMFGLTSL 201

Qy         152 EDLDLSYN-------------------NLH-----GLPWDSVRRMVNLHQLSLDHNLLDH 187
               | ||||||                   :||      ||  | | :  | :| |  | :|
Db         202 EVLDLSYNQIQSFHISSWSHTPKLKWLSLHSNRIQSLPSGSFRVLRQLEELILSANSIDS 261

Qy         188 IAEGTFADLQKLARLDLTSNRLQ---------------------------KLPPDPIFAR 220
               : :     :  | :|||:|| |                            :: |   | |
Db         262 LHKFALVGMSSLHKLDLSSNTLAVCVEDGAVLYNTSMPFLRSLRFTNNQLRVIPKRAFER 321

Qy         221 SQASA---LTATPFA-------PPLSFS---FGGNPLHCNCELLWLR------RLERDDD 261
                 |     ||  | |        ||         : : |:|:: ||       :|::
Db         322 FPALEELDLTDNPIATIHPEAFEPLELKRLVMNSSSILCDCQISWLASWIYRLKLDKSSI 381

Qy         262 LETCGSPGGLKGRYFWHVREEEFVC--EPP--LITQHTHKLLVLEGQAATLKCKAIGDPS 317
               :  |  |  |   |   :      |  : |   | :   ::  | |: |   |   |  |
Db         382 IAKCSYPPPLADLYVVAIDTANLTCHNDSPRAKIVRQPVEVSTLIGEKARFTCNVYG-AS 440

Qy         318 PL-IHW----------VAPDDRLVGNSSRTAVYDNGTLD----------IFITTSQDSGA 356
               || | |          :  |     : :||||  ||| |          :      |:
Db         441 PLSIEWRVMENGQPRVLVQDSATFLSINRTAVV-NGTFDERELAAAELLLDNVAMTDNSE 499

Qy         357 FTCIAANAAG-EATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGG 415
               : |:| |  | : :  |:: : | |  :
Db         500 YQCVARNRFGSDFSTHVKLQVYQAPKFT-------------------------------- 527

Qy         416 EPPKSPPERAVLVSEVT--------TTSALVKWSVSK-SAPRVKMYQLQYNCSDDEVLIY 466
                   :| :  :|| :          |    :||:  :   |  :  :|    :|| :
Db         528 ---YTPEDMPLLVGQTAKFLCAATGTPRPEIKWAFEQIPFPAAEARRLYVTPNDDHI--- 581

Qy         467 RMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQ----------FFTK 516
                        :::|            :   |  | |  |||: |  |          ||
Db         582 ---------YIMN------------VTKEDQGAYTCHATNVAGQTQASANLIVFENFF-- 618

Qy         517 ADYPQCQSMHSQIL 530
                 ||:   :   ::
Db         619 -HYPESPDLSPMLI 631

RESULT 10
JC1282
insulin-like growth factor-binding protein acid labile chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
C;Accession: JC1282
R;Dai, J.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 188, 304-309, 1992
A;Title: Molecular cloning of the acid-labile subunit of the rat insulin-like growth fac
A;Reference number: JC1282; MUID:93038676; PMID:1384485
A;Accession: JC1282
A;Molecule type: mRNA
A;Residues: 1-603 <DAI>
```

```
A;Cross-references: GB:S46785; NID:g258002; PIDN:AAB23770.2; PID:g5705934
A;Experimental source: liver
A;Note: the authors translated the codon AAG for residue 63 as Arg, AAA for residue 205
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-603/Product: insulin-like growth factor binding protein, acid labile chain #status
F;267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

  Query Match             6.1%;  Score 248;  DB 2;  Length 603;
  Best Local Similarity   25.2%;  Pred. No. 1.5e-08;
  Matches  104;  Conservative   64;  Mismatches  161;  Indels   84;  Gaps   15;

Qy          8 LLAFGMAF-----------AVVDA----CPKYCVC--QNLSESLGTLCPSKGLLFVPPDI 50
              |||| :|            |  ||    ||  | |   : :: |   | || |  :| ||
Db         13 LLAFWVALGPCHLQGTDPGASADAEGPQCPVACTCSHDDYTDELSVFCSSKNLTHLPDDI 72

Qy         51 DRRTVELRLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSN 110
                 |  | | || :  |    | |:: |  | |  : :  ::| : | |::|  || : |
Db         73 PVSTRALWLDGNNLSSIPSAAFQNLSSLDFLNLQGSWLRSLEPQALLGLQNLYYLHLERN 132

Qy        111 RLPSLGEDTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVR 170
              || :|         :|  | :::| || : :  |:  |  | ||:| :|:| .||    :
Db        133 RLRNLAVGLFTHTPSLASLSLSSNLLGRLEEGLFQG-LSHLWDLNLGWNSLVVLPDTVFQ 191

Qy        171 RMVNLHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATP 230
               : |||:| |  | | ::    |  | :|  |||: | |: :         :|:     |
Db        192 GLGNLHELVLAGNKLTYLQPALFCGLGELRELDLSRNALRSV---------KANVFVHLP 242

Qy        231 FAPPLSFSFGGNPLHCNCELLWLRRLERDDDLETCGSPGGLKG--RYFW----HVR---- 280
                                    |::|  | :| |  :||   |     |    | |
Db        243 ---------------------RLQKLYLDRNLITAVAPGAFLGMKALRWLDLSHNRVAGL 281

Qy        281 -EEEFVCEPPLITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVY 339
               |: |   | |:  |   | :     |:|: :   |    :|::  ::  :|::    :
Db        282 MEDTF---PGLLGLHV--LRLAHNAIASLRPRTFKD----LHFL--EELQLGHNRIRQLG 330

Qy        340 DN-----GTLDIF---------ITTSQDSGAFTCIAANAAGEATAMVEVSIVQ 378
              :      | |::          :     || |     | :|     :   : |
Db        331 ERTFEGLGQLEVLTLNDNQITEVRVGAFSGLFNVAVMNLSGNCLRSLPERVFQ 383


RESULT 11
A41915
insulin-like growth factor-binding complex acid-labile chain precursor - human
N;Alternate names: Acid-Labile Subunit (ALS)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C;Accession: A41915
R;Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.
Mol. Endocrinol. 6, 870-876, 1992
A;Title: Structure and functional expression of the acid-labile subunit of the insulin-l
A;Reference number: A41915; MUID:92357025; PMID:1379671
A;Accession: A41915
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-605 <LEO>
A;Cross-references: GB:M86826; NID:g184807; PIDN:AAA36047.1; PID:g184808
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIP:110171)
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
F;75-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;99-122/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;123-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;147-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;171-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
F;315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
F;339-362/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F;363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
```

```
F;387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F;411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F;435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
F;459-482/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
F;483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
F;507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR19>

  Query Match             5.9%;  Score 241.5;  DB 2;  Length 605;
  Best Local Similarity   29.7%;  Pred. No. 3.9e-08;
  Matches   76;  Conservative   41;  Mismatches  106;  Indels   33;  Gaps    4;

Qy         20 ACPKYCVC--QNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTG 77
              |||  |||   : :: |   | |: |  :|  :   |  | | || :  :    | |::
Db         40 ACPAACVCSYDDDADELSVFCSSRNLTRLPDGVPGGTQALWLDGNNLSSVPPAAFQNLSS 99

Qy         78 LVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLG 137
              |  | |    :  ::| : | ||:|  |||: |:| ||   |      |  | ::||:|
Db        100 LGFLNLQGGQLGSLEPQALLGLENLCHLHLERNQLRSLALGTFAHTPALASLGLSNNRLS 159

Qy        138 GIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQ 197
               : |  ||  | :| ||:| :|:|  ||  : | : :| :| |  | | ::    |: |
Db        160 RLEDGLFEG-LGSLWDLNLGWNSLAVLPDAAFRGLGSLRELVLAGNRLAYLQPALFSGLA 218

Qy        198 KLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLE 257
              :|  |||: | |:              |: |  |                 :|  |::|
Db        219 ELRELDLSRNALR--------------AIKANVFV-----------------QLPRLQKLY 248

Qy        258 RDDDLETCGSPGGLKG 273
               | :|    :||   |
Db        249 LDRNLIAAVAPGAFLG 264


RESULT 12
T42626
secreted leucine-rich repeat-containing protein SLIT2 - mouse (fragment)
N;Alternate names: neurogenic extracellular slit protein
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 16-Aug-2002
C;Accession: T42626
R;Holmes, G.P.; Negus, K.; Burridge, L.; Raman, S.; Algar, E.; Yamada, T.; Little, M.H.
Mech. Dev. 79, 57-72, 1998
A;Title: Distinct but overlapping expression patterns of two vertebrate slit homologs im
A;Reference number: Z22177; MUID:99279238; PMID:10349621
A;Accession: T42626
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1025 <HOL>
A;Cross-references: EMBL:AF074960; NID:g4151258; PID:g4151259; PIDN:AAD04345.1
C;Genetics:
A;Gene: Slit2
C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r

  Query Match             5.8%;  Score 237;  DB 2;  Length 1025;
  Best Local Similarity   20.4%;  Pred. No. 1.6e-07;
  Matches   86;  Conservative   60;  Mismatches  110;  Indels  166;  Gaps   10;

Qy         20 ACPKYCVCQNLSESLGTL--CPSKGLLFVPPDIDRRTVELRL------------------ 59
              |||: | |:      ||   | :: |  :|  | : | ||||
Db          1 ACPEKCRCE------GTTVDCSNQRLNKIPDHIPQYTAELRLNNNEFTVLEATGIFKKLP 54

Qy         60 -------GGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRL 112
                       | |  |    |   :|: :: |: | : ::|   |  ||||::| | |||:
Db         55 QLRXINFSNNKITDIEEGAFEGASGVNEILLTSNRLENVQHKMFKGLESLKTLMLRSNRI 114

Qy        113 PSLGEDTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYN----NLH------ 162
                :| |:  || ::: | : :||:  :|  || | | :|  |:|  |    | |
Db        115 SCVGNDSFIGLGSVRLLSLYDNQITTVAPGAF-DXLHSLSTLNLLANPFNCNCHLAWLGE 173

Qy        163 ------------------------------------------------------------ 162

Db        174 WLRRKRIVTGNPRCQKPYFLKEIPIQDVAIQDFTCDDGNDDNSCSPLSRCPSECTCLDTX 233

Qy        163 -------------GLPWDSVRRMVNLHQ----------------LSLDHNLLDHIAEGTF 193
```

```
                           |:| |     :: :|                : | :| :  ::   |
Db        234 VRCSNKGLKVLPKGIPKDVTELYLDGNQFTLVPKELSNYKHLTLIDLSNNRISTLSNQXF 293

Qy        194 ADLQKLARLDLTSNRLQKLPPD---------------------PIFARSQASALTATPFA 232
              ::: :|  | |: |||: :||                      |  | :  |||:
Db        294 SNMTQLLTLILSYNRLRCIPPRTFDGLKSLRLLSLHGNDISVVPEGAFNDLSALS----- 348

Qy        233 PPLSFSFGGNPLHCNCELLWL----RRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEP 288
                   : | |||:|:| : ||    :   ::  :  |  || :  :       ::| |:
Db        349 ---HLAIGANPLYCDCNMQWLSDWVKSEYKEPGIARCAGPGEMADKLLLTTPSKKFTCQG 405

Qy        289 PL 290
              |:
Db        406 PM 407


RESULT 13
A53860
chondroadherin precursor - bovine
N;Alternate names: 38K leucine-rich protein
C;Species: Bos primigenius taurus (cattle)
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999
C;Accession: A53860
R;Neame, P.J.; Sommarin, Y.; Boynton, R.E.; Heinegard, D.
J. Biol. Chem. 269, 21547-21554, 1994
A;Title: The structure of a 38-kDa leucine-rich protein (chondroadherin) isolated from b
A;Reference number: A53860; MUID:94342341; PMID:8063792
A;Accession: A53860
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-361 <NEA>
A;Cross-references: GB:U08018; NID:g470671; PIDN:AAA21330.1; PID:g470672
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan carboxyl-
C;Keywords: disulfide bond
F;300-346/Domain: proteoglycan carboxyl-terminal homology <PCH>

  Query Match             5.8%;  Score 236;  DB 2;  Length 361;
  Best Local Similarity   25.4%;  Pred. No. 4.4e-08;
  Matches   90;  Conservative   42;  Mismatches  122;  Indels  100;  Gaps   11;

Qy          7 GLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIH 66
              ||||       : |||: | |   |:    :|   ||  : | :  :|  | |  |
Db         13 GLLA--SLLPALAACPQNCHCH--SDLQHVICDKVGLQKI-PKVSEKTKLLNLQRNNFPV 67

Qy         67 ISRQDFANMTGLVD-------------------------LTLSRNTISHIQPFSFLDLESL 102
              ::   |  |  ||                         | || | |  ::  :| ||  |
Db         68 LATNSFRAMPNLVSLHLQHCQIREVAAGAFRGLKQLIYLYLSHNDIRVLRAGAFDDLTEL 127

Qy        103 RSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADEAFE----------------- 145
                |:|| |::  |    |  ||||  | :|||::  :   ||:
Db        128 TYLYLDHNKVTELPRGLLSPLVNLFILQLNNNKIRELRSGAFQGAKDLRWLYLSENSLSS 187

Qy        146 ------------------------------DFLLTLEDLDLSYNNLHGLPWDSVRRMVN- 174
                                              |  :|:| ||:| |  :| :: :
Db        188 LQPGALDDVENLAKFYLDRNQLSSYPSAALSKLRVVEELKLSHNPLKSIPDNAFQSFGRY 247

Qy        175 LHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPP 234
              |  | ||:  |:  ::| |  :  |  : | :||| :|| :  |       ||:|
Db        248 LETLWLDNTNLEKFSDGAFLGVTTLKHVHLENNRLHQLPSN--FPFDSLETLTLT----- 300

Qy        235 LSFSFGGNPLHCNCELLWLRR-----LERDDDLETCGSPGGLKGRYFWHVREEE 283
                     ||  | |:|  |||       | |   || ||   :|:   |:|: :
Db        301 ------NNPWKCTCQLRGLRRWLEAKTSRPD--ATCASPAKFRGQ---HIRDTD 343

RESULT 14
JC5239
insulin-like growth factor acid-labile chain - baboon
C;Species: Papio sp. (baboon)
C;Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
C;Accession: JC5239
R;Delhanty, P.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 227, 897-902, 1996
```

```
A;Title: The cloning and expression of the baboon acid-labile subunit of the insulin-lik
A;Reference number: JC5239; MUID:97040714; PMID:8886027
A;Contents: liver
A;Accession: JC5239
A;Molecule type: mRNA
A;Residues: 1-605 <DEL>
C;Comment: This factor is structurally related to proinsulin and have insuline-like meta

  Query Match             5.8%;  Score 235.5;  DB 2;  Length 605;
  Best Local Similarity   30.5%;  Pred. No. 9.8e-08;
  Matches   78;  Conservative   34;  Mismatches  111;  Indels   33;  Gaps    4;

Qy         20 ACPKYCVCQNLSE--SLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTG 77
              |||  | |    |   |   | |: |  :|  |   |  | |  | :  |    |  |::
Db         40 ACPATCACSYDDEVNELSVFCSSRNLTRLPDGIPGGTQALWLDSNNLSSIPPAAFRNLSS 99

Qy         78 LVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLG 137
              |  | |    :  ::| : | ||:|  |||: |:| ||   |      |  | ::||:|
Db        100 LAFLNLQGGQLGSLEPQALLGLENLCHLHLERNQLRSLAVGTFAYTPALALLGLSNNRLS 159

Qy        138 GIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQ 197
               : |  ||  |  | ||:| :|:|  ||  : | :  | :| |  | | ::    |: |
Db        160 RLEDGLFEG-LGNLWDLNLGWNSLAVLPDAAFRGLGGLRELVLAGNRLAYLQPALFSGLA 218

Qy        198 KLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLE 257
              :|  |||: | |:              |: |  ||                ::|  |::|
Db        219 ELRELDLSRNALR--------------AIKANVFA----------------QLPRLQKLY 248

Qy        258 RDDDLETCGSPGGLKG 273
               | :|    :||   |
Db        249 LDRNLIAAVAPGAFLG 264


RESULT 15
S46224
peroxidasin - fruit fly (Drosophila sp.)
C;Species: Drosophila sp.
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Oct-2000
C;Accession: S46224
R;Nelson, R.E.; Fessler, L.I.; Takagi, Y.; Blumberg, B.; Keene, D.R.; Olson, P.F.; Parke
EMBO J. 13, 3438-3447, 1994
A;Title: Peroxidasin: a novel enzyme-matrix protein of Drosophila development.
A;Reference number: S46224; MUID:94341255; PMID:8062820
A;Accession: S46224
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1535 <NEL>
A;Cross-references: GB:U11052; NID:g531384; PIDN:AAA61568.1; PID:g531385
C;Superfamily: peroxidasin; myeloperoxidase homology; proteoglycan amino-terminal homolo
F;19-44/Domain: proteoglycan amino-terminal homology <PAH4>
F;661-1350/Domain: myeloperoxidase homology <MPX>

  Query Match             5.7%;  Score 232.5;  DB 2;  Length 1535;
  Best Local Similarity   23.5%;  Pred. No. 5.5e-07;
  Matches  102;  Conservative   56;  Mismatches  193;  Indels   83;  Gaps   13;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVP--PDIDRRTVELR 58
              |  ||| ||        |  || | |   :      |    |  ||  |  | :|::||
Db          7 MLQLLGLLLLLAGGVQSV-YCPAGCTCLERT----VRCIRAKLSAVPKLPQ-DTQTLDLR 60

Qy         59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
                 | |  :    |: :  |  | |: | ::::|  :   | :|| ::|::|||  |
Db         61 F--NHIEELPANAFSGLAQLTTLFLNDNELAYLQDGALNGLTALRFVYLNNNRLSRLPAT 118

Qy        119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                : :  |: : : || :  :    | | |  |  | :  | |  || |   |: || :|
Db        119 IFQRMPRLEGIFLENNDIWQLPAGLF-DNLPRLNRLIMYNNKLTQLPVDGFNRLNNLKRL 177

Qy        179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
               || | :|
Db        178 RLDGNAID---------------------------------------------------- 185

Qy        239 FGGNPLHCNCEL--LWLRRLERDDDLE------TCGSPGGLKGRYFWHVREEEFVCEPPL 290
```

```
                  : ||| :  || ||   |   :      || :|  |: : |  : |  | |  |
Db       186 -----IDCNCGVYSLW-RRWHLDVQRQLVSISLTCAAPQMLQNQGFSSLGEHHFKCAKPQ 239

Qy       291 ITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRL-VGNSSRTAVYDNGTLDIFIT 349
                      |  |:   | |:  |   | | |:     | :   ::  :  :|:|
Db       240 FLVAPQDAQVAAGEQVELSCEVTGLHRPQITWMHNTQELGLEEQTQAEILPSGSLLHRSA 299

Qy       350 TSQDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKT-SRG 408
              : | | : ||| |  |   :      |:|     :    :|  :| : :   : | : |
Db       300 DTSDMGIYQCIARNEMGALRSQP----VRLVVNGGNHPLDSPIDARSNQVWADAGTPTHG 355

Qy       409 GGGSGGGEPPKSPP 422
                     |  |||
Db       356 ATPLPSPSPSHSPP 369

Search completed: June  8, 2003, 21:31:21
Job time : 50 secs
```

```
                        GenCore version 5.1.6
               Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         June  8, 2003, 19:06:53 ; Search time 36 Seconds
                                          (without alignments)
                                          909.023 Million cell updates/sec

Title:          US-09-831-846-2
Perfect score:  4094
Sequence:       1 METLLGGLLAFGMAFAVVDA..........DLVGARGTFGSSEWVMESTV 789

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters:      112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_40:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB ID                          Description
-------------------------------------------------------------------------------
    1   388.5    9.5    713  1  GAC1_HUMAN                  O75325 homo sapien
    2   282.5    6.9    646  1  FLR1_HUMAN                  Q9nzu1 homo sapien
    3     263    6.4   1480  1  SLIT_DROME                  P24014 drosophila
    4     248    6.1    603  1  ALS_RAT                     P35859 rattus norv
    5     244    6.0    359  1  CHAD_HUMAN                  O15335 homo sapien
    6   242.5    5.9    966  1  Y918_HUMAN                  O94991 homo sapien
    7   241.5    5.9    605  1  ALS_HUMAN                   P35858 homo sapien
    8   240.5    5.9    649  1  FLR3_HUMAN                  Q9nzu0 homo sapien
    9   240.5    5.9    951  1  LGR4_RAT                    Q9z2h4 rattus norv
   10     238    5.8    660  1  FLR2_HUMAN                  O43155 homo sapien
   11     236    5.8    361  1  CHAD_BOVIN                  Q27972 bos taurus
   12     236    5.8    977  1  Y848_HUMAN                  O94933 homo sapien
   13   235.5    5.8    605  1  ALS_PAPHA                   O02833 papio hamad
   14   234.5    5.7    951  1  LGR4_HUMAN                  Q9bxb1 homo sapien
   15   232.5    5.7    757  1  LGR7_HUMAN                  Q9hbx9 homo sapien
   16     231    5.6    603  1  ALS_MOUSE                   P70389 mus musculu
   17     226    5.5    358  1  CHAD_MOUSE                  O55226 mus musculu
   18     225    5.5    358  1  CHAD_RAT                    O70210 rattus norv
   19     224    5.5    373  1  ASPN_MOUSE                  Q99mq4 mus musculu
   20     221    5.4    907  1  LGR5_HUMAN                  O75473 homo sapien
   21     220    5.4    379  1  ASPN_HUMAN                  Q9bxn1 homo sapien
   22     211    5.2    368  1  PGS1_XENLA                  Q9ib75 xenopus lae
   23     211    5.2    567  1  GPV_RAT                     O08770 rattus norv
   24     208    5.1    907  1  LGR5_MOUSE                  Q9z1p4 mus musculu
   25     208    5.1   2012  1  DSCA_HUMAN                  O60469 homo sapien
   26     203    5.0    821  1  TRKB_MOUSE                  P15209 mus musculu
   27     201    4.9    626  1  GPBA_HUMAN                  P07359 homo sapien
   28     201    4.9    905  1  TLR3_MOUSE                  Q99mb1 mus musculu
   29   200.5    4.9    567  1  GPV_MOUSE                   O08742 mus musculu
   30     198    4.8    536  1  CBP8_HUMAN                  P22792 homo sapien
   31     197    4.8    560  1  GPV_HUMAN                   P40197 homo sapien
   32     196    4.8    821  1  TRKB_RAT                    Q63604 rattus norv
   33     195    4.8    360  1  PGS2_BOVIN                  P21793 bos taurus
```

```
   34     195    4.8   1097  1  TOLL_DROME                  P08953 drosophila
   35   194.5    4.8    360  1  PGS2_PIG                    Q9xsd9 sus scrofa
   36   194.5    4.8   1041  1  TLR8_HUMAN                  Q9nr97 homo sapien
   37     194    4.7    360  1  PGS2_SHEEP                  Q9tte2 ovis aries
   38   191.5    4.7    360  1  PGS2_HORSE                  O46542 equus cabal
   39     190    4.6    737  1  LGR8_MOUSE                  Q91zz5 mus musculu
   40   189.5    4.6    338  1  LUM_MOUSE                   P51885 mus musculu
   41   186.5    4.6    375  1  FMOD_BOVIN                  P13605 bos taurus
   42   186.5    4.6   1315  1  CHAO_DROME                  P12024 drosophila
   43     186    4.5    360  1  PGS2_CANFA                  Q29393 canis famil
   44   185.5    4.5    354  1  PGS2_MOUSE                  P28654 mus musculu
   45   185.5    4.5    369  1  PGS1_BOVIN                  P21809 bos taurus

                                 ALIGNMENTS

RESULT 1
GAC1_HUMAN
ID   GAC1_HUMAN     STANDARD;      PRT;   713 AA.
AC   O75325;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glioma amplified on chromosome 1 protein precursor.
GN   GAC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Glial tumor;
RX   MEDLINE=98324709; PubMed=9662332;
RA   Malfoy B., Almeida A., Zhu X.X., Vogt N., Tyagi R., Muleris M.,
RA   Dutrillaux A.-M., Dutrillaux B., Ross D., Hanash S.;
RT   "GAC1, a new member of the leucine-rich repeat superfamily on
RT   chromosome band 1q32.1, is amplified and overexpressed in malignant
RT   gliomas.";
RL   Oncogene 16:2997-3002(1998).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF030435; AAC39792.1; -.
DR   MIM; 605492; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 9.
DR   Pfam; PF01462; LRRNT; 1.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00370; LRR; 6.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 2.
```

```
KW   Immunoglobulin domain; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    713       GLIOMA AMPLIFIED ON CHROMOSOME 1 PROTEIN.
FT   DOMAIN       19    630       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    631    651       POTENTIAL.
FT   DOMAIN      652    713       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       92    115       LRR 1.
FT   REPEAT      116    139       LRR 2.
FT   REPEAT      140    163       LRR 3.
FT   REPEAT      165    187       LRR 4.
FT   REPEAT      188    211       LRR 5.
FT   REPEAT      213    235       LRR 6.
FT   REPEAT      236    259       LRR 7.
FT   REPEAT      261    283       LRR 8.
FT   REPEAT      309    333       LRR 9.
FT   REPEAT      334    357       LRR 10.
FT   REPEAT      359    385       LRR 11.
FT   DOMAIN      438    504       IG-LIKE C2-TYPE DOMAIN.
FT   DISULFID    445    497       BY SIMILARITY.
FT   CARBOHYD     94     94       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    381    381       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    555    555       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    583    583       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   713 AA;  78798 MW;  EC8BC0FD240C9396 CRC64;

  Query Match             9.5%;  Score 388.5;  DB 1;  Length 713;
  Best Local Similarity   22.5%;  Pred. No. 1e-17;
  Matches  170;  Conservative   91;  Mismatches  242;  Indels  251;  Gaps   27;

Qy         21 CPKYCVCQ---------NLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQD 71
              ||  | ||         :  |:    |    |  ||| :   |  | |  | |: : : :
Db         29 CPPQCACQIRPWYTPRSSYREATTVDCNDLFLTAVPPALPAGTQTLLLQSNSIVRVDQSE 88

Qy         72 FANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIV 131
                 :  | :| ||:|: |  :   |  |  | ||||: |:|  | : :  || :|| | :
Db         89 LGYLANLTELDLSQNSFSDARDCDFHALPQLLSLHLEENQLTRLEDHSFAGLASLQELYL 148

Qy        132 NNNQLGGIADEAF----------------------------------------------- 144
              |:|||  ||  ||
Db        149 NHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFEMLPNLEILMIGGNKVDAILDMN 208

Qy        145 --------------------EDFLL----TLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
                                   |: |    :|| |    | |  :|  :: ::  |  | |
Db        209 FRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQLARVPRRALEQVPGLKFLDL 268

Qy        181 DHNLLDHIAEGTFA-------------------------DLQKLARLDLTSN-RLQKLPP 214
              : | |  :  | ||                         :| :| :||:|:| ||  : |
Db        269 NKNPLQRVGPGDFANMLHLKELGLNNMEELVSIDKFALVNLPELTKLDITNNPRLSFIHP 328

Qy        215 DPIFARSQASALTATPFA------------PPL-SFSFGGNPLHCNCELLW-------LR 254
                     |   |     |            | |      |||: |:| : |       :|
Db        329 RAFHHLPQMETLMLNNNALSALHQQTVESLPNLQEVGLHGNPIRCDCVIRWANATGTRVR 388

Qy        255 RLERDDDLETCGSPGGLKGRYFWHVREEEF-----VCEPPLITQHTH--KLLVLEGQAAT 307
               :|    |  |  |  |:      |||  |      |  |||:  :    | |  |::
Db        389 FIEPQSTL--CAEPPDLQR---LPVREVPFREMTDHC-LPLISPRSFPPSLQVASGESMV 442

Qy        308 LKCKAIGDPSPLIHWVAPDD-RLVGNSS--RTAVYDNGTLDIFITTSQDSGAFTCIAANA 364
              | |:|: :| | |:|| |   ||    :  |  ||  |||::   |::::| :||:| |
Db        443 LHCRALAEPEPEIYWVTPAGLRLTPAHAGRRCRVYPEGTLELRRVTAEEAGLYTCVAQNL 502

Qy        365 AGEATAMVEVSI------------------VQLPHLSN-STSRTAPPKSRLSDITGSSKT 405
               |  |  | | :                  ||  |  :   |   || :  :::| || :
Db        503 VGADTKTVSVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNLTWSSAS 562

Qy        406 SRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLI 465
              |  | |:                      :|| :
Db        563 SLRGQGA----------------------TALAR-------------------------- 574

Qy        466 YRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSM 525
                 :|    :: :  |:  | |  |:   : |  | |       |    ||    :  |
Db        575 ---LPRGTHSYNITRLLQATEYWACLQVAFADAHTQL------ACVWARTK----EATSC 621

Qy        526 HSQILGGTMILVIGGIIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPP 585
              | : ||    |:         :: : :|::   :  |    :    |        | :|
Db        622 H-RALGDRPGLI--------AILALAVLLLAAGLAAHLGTGQPRKGV-------GGRRPL 665

Qy        586 PPS------SAPA----GAP---PQGPPKVVVRN 606
              ||:      |||:     ||   |  | : : |:
Db        666 PPAWAFWGWSAPSVRVVSAPLVLPWNPGRKLPRS 699

RESULT 2
FLR1_HUMAN
ID   FLR1_HUMAN     STANDARD;      PRT;   646 AA.
AC   Q9NZU1;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Leucine-rich repeat transmembrane protein FLRT1 precursor
DE   (Fibronectin-like domain-containing leucine-rich transmembrane protein
DE   1).
GN   FLRT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND GYLCOSYLATION.
RX   MEDLINE=20112755; PubMed=10644439;
RA   Lacy S.E., Bonnemann C.G., Buzney E.A., Kunkel L.M.;
RT   "Identification of FLRT1, FLRT2, and FLRT3: a novel family of
RT   transmembrane leucine-rich repeat proteins.";
RL   Genomics 62:417-426(1999).
CC   -!- FUNCTION: May have a function in cell adhesion and/or receptor
CC       signaling.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: Expressed in kidney and brain.
CC   -!- PTM: N-glycosylated.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF169675; AAF28459.1; ALT_INIT.
DR   Genew; HGNC:3760; FLRT1.
DR   MIM; 604806; -.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00560; LRR; 7.
DR   Pfam; PF01462; LRRNT; 1.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00370; LRR; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 1.
KW   Transmembrane; Leucine-rich repeat; Repeat; Glycoprotein; Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    646       LEUCINE-RICH REPEAT TRANSMEMBRANE PROTEIN
FT                                FLRT1.
FT   DOMAIN       21    524       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    525    545       POTENTIAL.
FT   DOMAIN      546    646       CYTOPLASMIC (POTENTIAL).
```

```
FT   REPEAT       52     77       LRR 1.
FT   REPEAT       78     98       LRR 2.
FT   REPEAT       99    121       LRR 3.
FT   REPEAT      123    147       LRR 4.
FT   REPEAT      148    169       LRR 5.
FT   REPEAT      170    192       LRR 6.
FT   REPEAT      194    218       LRR 7.
FT   REPEAT      219    241       LRR 8.
FT   REPEAT      242    264       LRR 9.
FT   REPEAT      265    288       LRR 10.
FT   DOMAIN      407    485       FIBRONECTIN TYPE-III.
FT   CARBOHYD    221    221       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   646 AA;  71359 MW;  FF2BF5DC3CA13C92 CRC64;

  Query Match             6.9%;  Score 282.5;  DB 1;  Length 646;
  Best Local Similarity   22.2%;  Pred. No. 6.4e-11;
  Matches  145;  Conservative   84;  Mismatches  218;  Indels  205;  Gaps   27;

Qy          7 GLLAFGMAFAVVDA--CPKYCVCQNLSESLGTL-CPSKGLLFVPPDI--DRRTV------ 55
              ||:||     |:|:  ||  | | |     | : |  :||  :| ||  |  |:
Db         12 GLIAF--LTEVIDSTTCPSVCRCDN-----GFIYCNDRGLTSIPADIPDDATTLYLQNNQ 64

Qy         56 --------------------------------------ELRLGGNFIIHISRQDFANMTG 77
                                                    || |  | :  |:|   | :
Db         65 INNAGIPQDLKTKVNVQVIYLYENDLDEFPINLPRSLRELHLQDNNVRTIARDSLARIPL 124

Qy         78 LVDLTLSRNTIS--HIQPFSFLDLESLR---------------------SLHLDSNRLPS 114
              |  | |  |::|   |:  :| | : |:                      | || ||: :
Db        125 LEKLHLDDNSVSTVSIEEDAFADSKQLKLLFLSRNHLSSIPSGLPHTLEELRLDDNRIST 184

Qy        115 LGEDTLRGLVNLQHLIVNNNQLGG--IADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRM 172
              :     :|| :|: |::: | |    |||: |   |  | :| |  |:|   | :
Db        185 IPLHAFKGLNSLRRLVLDGNLLANQRIADDTFSR-LQNLTELSLVRNSLAAPPLN--LPS 241

Qy        173 VNLHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPP---DPIFARSQASALTAT 229
               :| :| |  | : ||   | | :::| ||||::| |  ||    | :   :|
Db        242 AHLQKLYLQDNAISHIPYNTLAKMRELERLDLSNNNLTTLPRGLFDDLGNLAQ------- 294

Qy        230 PFAPPLSFSFGGNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPP 289
                          ||  | | |:|||                      |
Db        295 -------LLLRNNPWFCGCNLMWLRD---------------------W------------ 314

Qy        290 LITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYD-NGTLDIFI 348
                         :: :|| :  : :    |        ::: |     |: |    :|
Db        315 -----------VKARAAVVNVRGLMCQGP--------EKVRG----MAIKDITSEMDECF 351

Qy        349 TTSQDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAP-----------PKSRLS 397
               |    |      |||| : ||            ||  | | |           |  ||
Db        352 ETGPQGG-----VANAAAKTTA------------SNHASATTPQGSLFTLKAKRPGLRLP 394

Qy        398 DITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYN 457
              |       : | |              |: |  :|  |  : |  :  |   ::  |:
Db        395 DSNIDYPMATGDGAK----------TLAIHVKALTADSIRITWKATLPASSFRLSWLRLG 444

Qy        458 CSDDEVLIYRMIPASNKA-FVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTK 516
               |     |   :   :|  :::  |   : | :|:: |  :|:    |     ||:  |
Db        445 HSPAVGSITETLVQGDKTEYLLTALEPKSTYIICMVTM--ETSNAYVADETPVCAKAETA 502

Qy        517 ADYPQCQSMHSQILGGTMI-LVIGGII---VATLLVFIVILMVRYKVCNHEA 564
                |    ::: :   | |  | : |||   || : :|:|:  : : |  |:|
Db        503 DSYGPTTTLNQEQNAGPMASLPLAGIIGGAVALVFLFLVLGAICWYV--HQA 552


RESULT 3
SLIT_DROME
ID   SLIT_DROME     STANDARD;      PRT;  1480 AA.
AC   P24014;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Slit protein precursor.
GN   SLI.
```

```
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91099665; PubMed=2176636;
RA   Rothberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;
RT   "Slit: an extracellular protein necessary for development of midline
RT   glia and commissural axon pathways contains both EGF and LRR
RT   domains.";
RL   Genes Dev. 4:2169-2187(1990).
CC   -!- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND
CC       COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR
CC       MATRIX MOLECULES.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND
CC       EVENTUALLY DISTRIBUTED ALONG THE AXONS.
CC   -!- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X53959; CAA37910.1; -.
DR   PIR; A36665; A36665.
DR   HSSP; P00740; 1EDM.
DR   FlyBase; FBgn0003425; sli.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000359; Cys_knot.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   InterPro; IPR001791; Laminin_G.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00008; EGF; 7.
DR   Pfam; PF00054; laminin_G; 1.
DR   Pfam; PF00560; LRR; 16.
DR   Pfam; PF01462; LRRNT; 4.
DR   Pfam; PF01463; LRRCT; 4.
DR   SMART; SM00041; CT; 1.
DR   SMART; SM00179; EGF_CA; 2.
DR   SMART; SM00001; EGF_like; 5.
DR   SMART; SM00370; LRR; 4.
DR   SMART; SM00082; LRRCT; 4.
DR   SMART; SM00013; LRRNT; 4.
DR   SMART; SM00369; LRR_TYP; 9.
DR   SMART; SM00282; LamG; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00022; EGF_1; 7.
DR   PROSITE; PS01186; EGF_2; 5.
DR   PROSITE; PS01187; EGF_CA; 2.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 1.
KW   Neurogenesis; Glycoprotein; Signal; Alternative splicing;
KW   EGF-like domain; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     36
FT   CHAIN        37   1480       SLIT PROTEIN.
```

```
FT   REPEAT        99    122       LRR 1.
FT   REPEAT       123    146       LRR 2.
FT   REPEAT       148    170       LRR 3.
FT   REPEAT       171    194       LRR 4.
FT   REPEAT       195    218       LRR 5.
FT   REPEAT       220    246       LRR 6.
FT   REPEAT       321    344       LRR 7.
FT   REPEAT       345    368       LRR 8.
FT   REPEAT       369    392       LRR 9.
FT   REPEAT       394    416       LRR 10.
FT   REPEAT       417    440       LRR 11.
FT   REPEAT       522    545       LRR 12.
FT   REPEAT       546    569       LRR 13.
FT   REPEAT       570    593       LRR 14.
FT   REPEAT       595    617       LRR 15.
FT   REPEAT       618    641       LRR 16.
FT   REPEAT       643    666       LRR 17.
FT   REPEAT       678    701       LRR 18.
FT   REPEAT       720    743       LRR 19.
FT   REPEAT       745    764       LRR 20.
FT   REPEAT       765    788       LRR 21.
FT   REPEAT       790    812       LRR 22.
FT   REPEAT       813    836       LRR 23.
FT   REPEAT       838    861       LRR 24.
FT   DOMAIN       907    944       EGF-LIKE 1.
FT   DOMAIN       946    983       EGF-LIKE 2.
FT   DOMAIN       985   1022       EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      1024   1062       EGF-LIKE 4.
FT   DOMAIN      1064   1100       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      1111   1149       EGF-LIKE 6.
FT   DOMAIN      1152   1325       LAMININ G-LIKE.
FT   DOMAIN      1353   1392       EGF-LIKE 7.
FT   DOMAIN      1409   1480       CTCK.
FT   CARBOHYD     111    111       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     207    207       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     357    357       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     435    435       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     783    783       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     788    788       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     958    958       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     998    998       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1060   1060       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1159   1159       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1175   1175       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1243   1243       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1292   1292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     911    922       BY SIMILARITY.
FT   DISULFID     916    932       BY SIMILARITY.
FT   DISULFID     934    943       BY SIMILARITY.
FT   DISULFID     950    961       BY SIMILARITY.
FT   DISULFID     955    971       BY SIMILARITY.
FT   DISULFID     973    982       BY SIMILARITY.
FT   DISULFID     989   1001       BY SIMILARITY.
FT   DISULFID     995   1010       BY SIMILARITY.
FT   DISULFID    1012   1021       BY SIMILARITY.
FT   DISULFID    1028   1041       BY SIMILARITY.
FT   DISULFID    1035   1050       BY SIMILARITY.
FT   DISULFID    1052   1061       BY SIMILARITY.
FT   DISULFID    1068   1079       BY SIMILARITY.
FT   DISULFID    1073   1088       BY SIMILARITY.
FT   DISULFID    1090   1099       BY SIMILARITY.
FT   DISULFID    1115   1125       BY SIMILARITY.
FT   DISULFID    1120   1137       BY SIMILARITY.
FT   DISULFID    1139   1148       BY SIMILARITY.
FT   DISULFID    1357   1368       BY SIMILARITY.
FT   DISULFID    1362   1380       BY SIMILARITY.
FT   DISULFID    1382   1391       BY SIMILARITY.
FT   DISULFID    1409   1443       BY SIMILARITY.
FT   DISULFID    1423   1457       BY SIMILARITY.
FT   DISULFID    1434   1473       BY SIMILARITY.
FT   DISULFID    1438   1475       BY SIMILARITY.
FT   DISULFID    1442   1479       BY SIMILARITY.
FT   VARSPLIC    1394   1404       MISSING (IN SHORT ISOFORM).
```

```
SQ   SEQUENCE   1480 AA;  165752 MW;  F9D5925FC170B1C3 CRC64;

  Query Match             6.4%;  Score 263;  DB 1;  Length 1480;
  Best Local Similarity   28.0%;  Pred. No. 3.7e-09;
  Matches   89;  Conservative   39;  Mismatches  102;  Indels   88;  Gaps   11;

Qy          6 GGLLAFGMAF------AVVDA-CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELR 58
              ||| : |:         : :| ||: | |  |:      |  :||  ||  |      |
Db         51 GGLGSVGIHIPGGGVGVITEARCPRVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106

Qy         59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
              | || :  |   ||  :| |  | |: | |  |:  || || ||  | : :| : ::|
Db        107 LQGNNLTVIYETDFQRLTKLRMLQLTDNQIHTIERNSFQDLVSLERLDISNNVITTVGRR 166

Qy        119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                :|  :|: | ::|||:  : : ||:  |: || | |: |||  ||
Db        167 VFKGAQSLRSLQLDNNQITCLDEHAFKG-LVELEILTLNNNNLTSLP------------ 212

Qy        179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFS 238
                 ||:        |  :|  | |: |                      |||
Db        213 ---HNI--------FGGLGRLRALRLSDN----------------------PFA------ 233

Qy        239 FGGNPLHCNCELLWLRRLERD----DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQH 294
                     |:| | || |  |          | ||  |||:    : ::|| |    :|:|
Db        234 -------CDCHLSWLSRFLRSATRLAPYTRCQSPSQLKGQNVADLHDQEFKCSG--LTEH 284

Qy        295 THKLLVLEGQAATLKCKA 312
                         | ::| |
Db        285 -----------APMECGA 291


RESULT 4
ALS_RAT
ID   ALS_RAT        STANDARD;      PRT;   603 AA.
AC   P35859;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Insulin-like growth factor binding protein complex acid labile chain
DE   precursor (ALS).
GN   IGFALS OR ALS.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93038676; PubMed=1384485;
RA   Dai J., Baxter R.C.;
RT   "Molecular cloning of the acid-labile subunit of the rat insulin-like
RT   growth factor binding protein complex.";
RL   Biochem. Biophys. Res. Commun. 188:304-309(1992).
RN   [2]
RP   SEQUENCE OF 24-44, AND CHARACTERIZATION.
RC   STRAIN=Wistar; TISSUE=Serum;
RX   MEDLINE=94130835; PubMed=7507839;
RA   Baxter R.C., Dai J.;
RT   "Purification and characterization of the acid-labile subunit of rat
RT   serum insulin-like growth factor binding protein complex.";
RL   Endocrinology 134:848-852(1994).
CC   -!- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF
CC       CIRCULATING IGFS TO THE TISSUES.
CC   -!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
CC       IGF-I OR IGF-II AND IGFBP-3.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: BRAIN, KIDNEY, LUNG, HEART, SPLEEN, MUSCLE
CC       AND LIVER.
CC   -!- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
```

```
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; S46785; AAB23770.2; -.
DR   PIR; JC1282; JC1282.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00560; LRR; 18.
DR   Pfam; PF01462; LRRNT; 1.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00370; LRR; 5.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 9.
KW   Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    603       INSULIN-LIKE GROWTH FACTOR BINDING
FT                                PROTEIN COMPLEX ACID LABILE CHAIN.
FT   REPEAT       52     73       LRR 1.
FT   REPEAT       74     96       LRR 2.
FT   REPEAT       98    120       LRR 3.
FT   REPEAT'     121    144       LRR 4.
FT   REPEAT      146    168       LRR 5.
FT   REPEAT      169    192       LRR 6.
FT   REPEAT      194    216       LRR 7.
FT   REPEAT      217    240       LRR 8.
FT   REPEAT      242    264       LRR 9.
FT   REPEAT      266    288       LRR 10.
FT   REPEAT      289    312       LRR 11.
FT   REPEAT      313    336       LRR 12.
FT   REPEAT      337    360       LRR 13.
FT   REPEAT      361    384       LRR 14.
FT   REPEAT      386    408       LRR 15.
FT   REPEAT      409    432       LRR 16.
FT   REPEAT      433    456       LRR 17.
FT   REPEAT      458    478       LRR 18.
FT   REPEAT      479    504       LRR 19.
FT   REPEAT      506    529       LRR 20.
FT   REPEAT      543    566       LRR 21.
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     85     85       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     96     96       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    515    515       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    578    578       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    586    586       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   603 AA;  66811 MW;  DCD7637D94A5037C CRC64;

  Query Match             6.1%;  Score 248;  DB 1;  Length 603;
  Best Local Similarity   25.2%;  Pred. No. 1e-08;
  Matches  104;  Conservative   64;  Mismatches  161;  Indels   84;  Gaps   15;

Qy           8 LLAFGMAF-----------AVVDA----CPKYCVC--QNLSESLGTLCPSKGLLFVPPDI 50
               |||| :|            |  ||    ||  | |   : :: |   | || |  :| ||
Db          13 LLAFWVALGPCHLQGTDPGASADAEGPQCPVACTCSHDDYTDELSVFCSSKNLTHLPDDI 72

Qy          51 DRRTVELRLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSN 110
                  |  | | || :  |    | |:: |  | |  : :  ::| : | |::|  |||: |
Db          73 PVSTRALWLDGNNLSSIPSAAFQNLSSLDFLNLQGSWLRSLEPQALLGLQNLYYLHLERN 132

Qy         111 RLPSLGEDTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVR 170
               || :|         :|  | :::| || : :  |:   | | ||:| :|:|  ||    :
Db         133 RLRNLAVGLFTHTPSLASLSLSSNLLGRLEEGLFQG-LSHLWDLNLGWNSLVVLPDTVFQ 191

Qy         171 RMVNLHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATP 230
                : |||:| |  | | ::    |  | :|  |||: | |: :         :|:     |
Db         192 GLGNLHELVLAGNKLTYLQPALFCGLGELRELDLSRNALRSV---------KANVFVHLP 242

Qy         231 FAPPLSFSFGGNPLHCNCELLWLRRLERDDDLETCGSPGGLKG--RYFW----HVR---- 280
                                     |::|  | :| |  :||   |     |    | |
Db         243 ---------------------RLQKLYLDRNLITAVAPGAFLGMKALRWLDLSHNRVAGL 281

Qy         281 -EEEFVCEPPLITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVY 339
                |: |   | |:  |   | :     |:|: :   |    :|::  ::  :|::    :
Db         282 MEDTF---PGLLGLHV--LRLAHNAIASLRPRTFKD----LHFL--EELQLGHNRIRQLG 330

Qy         340 DN-----GTLDIF---------ITTSQDSGAFTCIAANAAGEATAMVEVSIVQ 378
               :      | |::          :     || |     | :|     :   : |
Db         331 ERTFEGLGQLEVLTLNDNQITEVRVGAFSGLFNVAVMNLSGNCLRSLPERVFQ 383


RESULT 5
CHAD_HUMAN
ID   CHAD_HUMAN     STANDARD;      PRT;   359 AA.
AC   O15335; Q96RJ5;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Chondroadherin precursor (Cartilage leucine-rich protein).
GN   CHAD.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98008928; PubMed=9344663;
RA   Grover J., Chen X.-N., Korenberg J.R., Roughley P.J.;
RT   "The structure and chromosome location of the human chondroadherin
RT   gene (CHAD).";
RL   Genomics 45:379-385(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21413956; PubMed=11445564;
RA   Maansson B., Wenglen C., Moergelin M., Saxne T., Heinegaard D.;
RT   "Association of chondroadherin with collagen type II.";
RL   J. Biol. Chem. 276:32883-32888(2001).
CC   -!- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and
CC       osteoblasts. This binding is mediated (at least for chondrocytes
CC       and fibroblasts) by the integrin alpha(2)beta(1). May play an
CC       important role in the regulation of chondrocyte growth and
CC       proliferation (By similarity).
CC   -!- SUBUNIT: Mostly monomeric (By similarity). Interacts with collagen
CC       type II.
CC   -!- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).
CC   -!- TISSUE SPECIFICITY: Present in chondrocytes at all ages.
CC   -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
CC       (SLRPS) FAMILY. CLASS IV SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U96769; AAC13410.1; -.
DR   EMBL; U96767; AAC13410.1; JOINED.
DR   EMBL; U96768; AAC13410.1; JOINED.
DR   EMBL; AF371328; AAK51556.1; -.
DR   Genew; HGNC:1909; CHAD.
DR   MIM; 602178; -.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00560; LRR; 10.
```

```
DR   Pfam; PF01463; LRRCT; 1.
DR   Pfam; PF01462; LRRNT; 1.
DR   SMART; SM00370; LRR; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 4.
KW   Signal; Leucine-rich repeat; Repeat; Glycoprotein;
KW   Extracellular matrix.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    359       CHONDROADHERIN.
FT   REPEAT       50     73       LRR 1.
FT   REPEAT       74     97       LRR 2.
FT   REPEAT       99    121       LRR 3.
FT   REPEAT      122    145       LRR 4.
FT   REPEAT      147    169       LRR 5.
FT   REPEAT      171    193       LRR 6.
FT   REPEAT      194    217       LRR 7.
FT   REPEAT      218    241       LRR 8.
FT   REPEAT      243    266       LRR 9.
FT   REPEAT      267    290       LRR 10.
FT   REPEAT      292    317       LRR 11.
FT   DISULFID     23     38       BY SIMILARITY.
FT   DISULFID    304    346       BY SIMILARITY.
FT   DISULFID    306    326       BY SIMILARITY.
FT   CARBOHYD    144    144       O-LINKED (POTENTIAL).
FT   CONFLICT    114    114       V -> L (IN REF. 2).
FT   CONFLICT    166    166       P -> A (IN REF. 2).
SQ   SEQUENCE   359 AA;  40487 MW;  9A318D0A15C157A0 CRC64;

  Query Match             6.0%;  Score 244;  DB 1;  Length 359;
  Best Local Similarity   26.1%;  Pred. No. 8.9e-09;
  Matches   93;  Conservative   41;  Mismatches  119;  Indels  104;  Gaps   11;

Qy           4 LLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNF 63
               || |||        : |||: | |   |:    :|   ||  : | :  :|  | |  |
Db          12 LLAGLL------PALAACPQNCHCH--SDLQHVICDKVGLQKI-PKVSEKTKLLNLQRNN 62

Qy          64 IIHISRQDFANMTGLVD-------------------------LTLSRNTISHIQPFSFLDL 99
                  ::   |  |  ||                         | || | |  ::  :| ||
Db          63 FPVLAANSFRAMPNLVSLHLQHCQIREVAAGAFRGLKQLIYLYLSHNDIRVVRAGAFDDL 122

Qy         100 ESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ------------------------ 135
                 |  |:|| |::  |    |  ||||  | :|||:
Db         123 TELTYLYLDHNKVTELPRGLLSPLVNLFILQLNNNKIRELRAGPFQGAKDLRWLYLSENA 182

Qy         136 LGGIADEAFEDF------------------------LLTLEDLDLSYNNLHGLPWDSVRRM 172
               |  :   | :|                        |  :|:| ||:| |  :| :: :
Db         183 LSSLQPGALDDVENLAKFHVDRNQLSSYPSAALSKLRVVEELKLSHNPLKSIPDNAFQSF 242

Qy         173 VN-LHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPF 231
                  |  | ||:  |:  ::| |  :  |  : | :||| :|| :  |   :  |||
Db         243 GRYLETLWLDNTNLEKFSDGAFLGVTTLKHVHLENNRLNQLPSNFPFDSLETLALT---- 298

Qy         232 APPLSFSFGGNPLHCNCELLWLRR-----LERDDDLETCGSPGGLKGRYFWHVREEE 283
                         ||  | |:|  |||       | |   || ||   ||:   |:|: :
Db         299 ---------NNPWKCTCQLRGLRRWLEAKASRPD--ATCASPAKFKGQ---HIRDTD 341


RESULT 6
Y918_HUMAN
ID    Y918_HUMAN      STANDARD;      PRT;   966 AA.
AC    O94991;
DT    16-OCT-2001 (Rel. 40, Created)
DT    16-OCT-2001 (Rel. 40, Last sequence update)
DT    16-OCT-2001 (Rel. 40, Last annotation update)
DE    Hypothetical protein KIAA0918 (Fragment).
GN    KIAA0918.
OS    Homo sapiens (Human).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX    NCBI_TaxID=9606;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Brain;
RX    MEDLINE=99156230; PubMed=10048485;
RA    Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA    Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT    "Prediction of the coding sequences of unidentified human genes. XII.
RT    The complete sequences of 100 new cDNA clones from brain which code
RT    for large proteins in vitro.";
RL    DNA Res. 5:355-364(1998).
CC    -!- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
CC    --------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC    the European Bioinformatics Institute.  There are no  restrictions on  its
CC    use  by  non-profit  institutions as long  as its content  is  in  no  way
CC    modified and this statement is not removed.  Usage  by  and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------------
DR    EMBL; AB020725; BAA74941.1; -.
DR    InterPro; IPR001611; LRR.
DR    InterPro; IPR000483; LRR_Cterm.
DR    InterPro; IPR000372; LRR_Nterm.
DR    InterPro; IPR003592; LRR_out.
DR    InterPro; IPR003591; LRR_typ.
DR    Pfam; PF00560; LRR; 9.
DR    Pfam; PF01463; LRRCT; 2.
DR    PRINTS; PR00019; LEURICHRPT.
DR    SMART; SM00370; LRR; 2.
DR    SMART; SM00082; LRRCT; 2.
DR    SMART; SM00013; LRRNT; 1.
DR    SMART; SM00369; LRR_TYP; 4.
KW    Hypothetical protein; Repeat; Leucine-rich repeat.
FT    NON_TER       1      1
FT    REPEAT       88    111       LRR 1.
FT    REPEAT      112    135       LRR 2.
FT    REPEAT      136    159       LRR 3.
FT    REPEAT      161    183       LRR 4.
FT    REPEAT      184    207       LRR 5.
FT    REPEAT      209    231       LRR 6.
FT    REPEAT      440    463       LRR 7.
FT    REPEAT      464    487       LRR 8.
FT    REPEAT      489    511       LRR 9.
FT    REPEAT      512    535       LRR 10.
FT    REPEAT      537    559       LRR 11.
SQ    SEQUENCE   966 AA;  108296 MW;  A1D236413E2B27DF CRC64;

  Query Match             5.9%;  Score 242.5;  DB 1;  Length 966;
  Best Local Similarity   21.2%;  Pred. No. 4.4e-08;
  Matches  177;  Conservative   96;  Mismatches  327;  Indels  233;  Gaps   34;

Qy          13 MAFAVV--------------DACPKYCVCQNLSESLGTLCPSKGLLFV----PPDIDRRT 54
               :||||               : |   | |:     |   | ::|:: :    ||
Db          34 LAFAVTSLVLSCAETIDYYGEICDNACPCEEKDGILTVSCENRGIISLSEISPPRFP--I 91

Qy          55 VELRLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPS 114
                 | | ||    :   :| | ||   | |  | |  |:  :|  |  || |||::|:|
Db          92 YHLLLSGNLLNRLYPNEFVNYTGASILHLGSNVIQDIETGAFHGLRGLRRLHLNNNKLEL 151

Qy         115 LGEDTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVN 174
               | :||  || ||::| |: | :  |   ||                ||
Db         152 LRDDTFLGLENLEYLQVDYNYISVIEPNAF---------------GKLH-----------L 186

Qy         175 LHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPP 234
               |  | |: |||  :    |     |  |||  |||: ||   :
Db         187 LQVLILNDNLLSSLPNNLFR-FVPLTHLDLRGNRLKLLPYVGLLQHMD----------KV 235

Qy         235 LSFSFGGNPLHCNCELL----WLRRLERDDDLE--------TCGSPGGLKGRYFWHVREE 282
               :      || :|:|||:    ||      | :          | :|  | ||    | ::
Db         236 VELQLEENPWNCSCELISLKDWL------DSISYSALVGDVVCETPFRLHGRDLDEVSKQ 289

Qy         283 EFVCEPPLI--------------TQHTHKLLVLEGQAATLKCKAIGDP--------SPLIH 321
               | :|   ||             | : |         ||        |         |
Db         290 E-LCPRRLISDYEMRPQTPLSTTGYLHTTPASVNSVATSSSAVYKPPLKPPKGTRQPNKP 348
```

```
Qy         322 WVAPDDRLVGNSSRTAVYDN-GTLDIFITTS----QDSGAFTC----------------- 359
                | |  |     |:   | | |    : | |    :   | :|
Db         349 RVRPTSR---QPSKDLGYSNYGPSIAYQTKSPVPLECPTACSCNLQISDLGLNVNCQERK 405

Qy         360 ------------------IAAN--AAGEATAMVEVSIVQLPHLSNSTSRTAPPKSR-LSD 398
                                 :  |  |    |  :| : : | || |  :| :  : |   |
Db         406 IESIAELQPKPYNPKKMYLTENYIAVVRRTDFLEATGLDLLHLGN--NRISMIQDRAFGD 463

Qy         399 ITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPR-------VKM 451
               :|   :    |                                : : :|        ::
Db         464 LTNLRRLYLNGN-------------------------------RIERLSPELFYGLQSLQY 493

Qy         452 YQLQYN-CSDDEVLIYRMIPASNKAFVVNNLV----SG--TGYDLCVLAMWDDTATTLTA 504
                 ||||   : :   :  :|     |: |||:    ||  :|  |  | :  :  |:|
Db         494 LFLQYNLIREIQSGTFDPVPNLQLLFLNNNLLQAMPSGVFSGLTLLRLNLRSNHFTSLPV 553

Qy         505 TNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRYKVCNHEA 564
               : ::   :   : |      :|      |  :|   : |  | | :::  |   :|  :|
Db         554 SGVLDQLKSLIQID------LHDNPWDCTCDIVGMKLWVEQLKVGVLVDEV---IC--KA 602

Qy         565 PSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPPKVVVRNELLDFTASLARASDSSSS 624
               | | |   ::: |  :    |  |      |     :|  |   :       |  : : :
Db         603 PKKFAE--TDMRSIKSELLCPDYSDVVVSTPTPSSIQVPARTSAV---TPAVRLNSTGAP 657

Qy         625 SSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAFAS-----LDLKSQRKEELLDSRT- 678
               :|||:|  |        :| |       |  :|  | :     | :|  ::| :   : |
Db         658 ASLGAGGGAS------SVPLSVLILSLLLVFIMSVFVAAGLFVLVMKRRKKNQSDHTSTN 711

Qy         679 ---------------PAGRGAGTSARGHHSDREPLL---GPPAARARSLLPLPL 714
                               | | |     |   | | |     ||      :|  ||
Db         712 NSDVSSFNMQYSVYGGGGGTGGHPHAHVHHRGPALPKVKTPAGHVYEYIPHPL 764

RESULT 7
ALS_HUMAN
ID   ALS_HUMAN      STANDARD;      PRT;   605 AA.
AC   P35858;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Insulin-like growth factor binding protein complex acid labile chain
DE   precursor (ALS).
GN   IGFALS OR ALS.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=92357025; PubMed=1379671;
RA   Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;
RT   "Structure and functional expression of the acid-labile subunit of
RT   the insulin-like growth factor-binding protein complex.";
RL   Mol. Endocrinol. 6:870-876(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Frankland J.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 28-35.
RX   MEDLINE=89308584; PubMed=2473065;
RA   Baxter R.C., Martin J.L., Beniac V.A.;
RT   "High molecular weight insulin-like growth factor binding protein
RT   complex. Purification and properties of the acid-labile subunit from
RT   human serum.";
RL   J. Biol. Chem. 264:11843-11848(1989).
CC   -!- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
CC       IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC   -!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
CC       IGF-I OR IGF-II AND IGFBP-3.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M86826; AAA36047.1; -.
DR   EMBL; AL031724; CAC36078.1; -.
DR   PIR; A41915; A41915.
DR   Genew; HGNC:5468; IGFALS.
DR   MIM; 601489; -.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00560; LRR; 19.
DR   Pfam; PF01462; LRRNT; 1.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00370; LRR; 2.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 11.
KW   Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT   SIGNAL        1     27
FT   CHAIN        28    605       INSULIN-LIKE GROWTH FACTOR BINDING
FT                                PROTEIN COMPLEX ACID LABILE CHAIN.
FT   REPEAT       53     73       LRR 1.
FT   REPEAT       74     96       LRR 2.
FT   REPEAT       98    120       LRR 3.
FT   REPEAT      121    144       LRR 4.
FT   REPEAT      145    168       LRR 5.
FT   REPEAT      170    192       LRR 6.
FT   REPEAT      193    216       LRR 7.
FT   REPEAT      217    240       LRR 8.
FT   REPEAT      242    264       LRR 9.
FT   REPEAT      265    288       LRR 10.
FT   REPEAT      289    312       LRR 11.
FT   REPEAT      313    336       LRR 12.
FT   REPEAT      337    360       LRR 13.
FT   REPEAT      361    384       LRR 14.
FT   REPEAT      386    408       LRR 15.
FT   REPEAT      409    432       LRR 16.
FT   REPEAT      433    456       LRR 17.
FT   REPEAT      458    480       LRR 18.
FT   REPEAT      482    504       LRR 19.
FT   REPEAT      505    530       LRR 20.
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     85     85       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     96     96       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    515    515       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    580    580       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   605 AA;  66034 MW;  F6562A23CBE918F6 CRC64;

  Query Match             5.9%;  Score 241.5;  DB 1;  Length 605;
  Best Local Similarity   29.7%;  Pred. No. 2.7e-08;
  Matches   76;  Conservative   41;  Mismatches  106;  Indels   33;  Gaps    4;

Qy          20 ACPKYCVC--QNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTG 77
               |||  |||   : :: |   | |: |  :|  :   |  | | || :  :    | |::
Db          40 ACPAACVCSYDDDADELSVFCSSRNLTRLPDGVPGGTQALWLDGNNLSSVPPAAFQNLSS 99

Qy          78 LVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLG 137
               |  | |    :  ::| : | ||:|  |||: |:| ||   |      |  | ::||:|
Db         100 LGFLNLQGGQLGSLEPQALLGLENLCHLHLERNQLRSLALGTFAHTPALASLGLSNNRLS 159
```

```
Qy        138 GIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQ 197
               : |  ||  | :| ||:| :|:|  ||  : | : :| :| |  | | ::    |: |
Db        160 RLEDGLFEG-LGSLWDLNLGWNSLAVLPDAAFRGLGSLRELVLAGNRLAYLQPALFSGLA 218

Qy        198 KLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLE 257
              :|  |||: | |:              |: |  |                 :|  |::|
Db        219 ELRELDLSRNALR--------------AIKANVFV----------------QLPRLQKLY 248

Qy        258 RDDDLETCGSPGGLKG 273
               | :|    :||   |
Db        249 LDRNLIAAVAPGAFLG 264


RESULT 8
FLR3_HUMAN
ID   FLR3_HUMAN     STANDARD;      PRT;   649 AA.
AC   Q9NZU0; Q96KB1; Q9P259; Q96K42; Q96K39;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Leucine-rich repeat transmembrane protein FLRT3 precursor
DE   (Fibronectin-like domain-containing leucine-rich transmembrane protein
DE   3).
GN   FLRT3 OR KIAA1469.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX   MEDLINE=20112755; PubMed=10644439;
RA   Lacy S.E., Bonnemann C.G., Buzney E.A., Kunkel L.M.;
RT   "Identification of FLRT1, FLRT2, and FLRT3: a novel family of
RT   transmembrane leucine-rich repeat proteins.";
RL   Genomics 62:417-426(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20277482; PubMed=10819331;
RA   Nagase T., Kikuno R., Ishikawa K.-I., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XVII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 7:143-150(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo, and Teratocarcinoma;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May have a function in cell adhesion and/or receptor
CC       signaling.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: Expressed in kidney, brain, pancreas, skeletal
CC       muscle, lung, liver, placenta, and heart.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF169677; AAF28461.1; -.
DR   EMBL; AB040902; BAA95993.1; ALT_INIT.
DR   EMBL; AK027297; BAB55023.1; -.
DR   EMBL; AK027670; BAB55282.1; ALT_INIT.
DR   EMBL; AK027694; BAB55303.1; ALT_INIT.
DR   EMBL; AL132826; CAB86687.1; -.
DR   EMBL; BC020870; AAH20870.1; -.
DR   Genew; HGNC:3762; FLRT3.
DR   MIM; 604808; -.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00560; LRR; 10.
DR   Pfam; PF01462; LRRNT; 1.
DR   Pfam; PF01463; LRRCT; 1.
DR   SMART; SM00060; FN3; 1.
DR   SMART; SM00370; LRR; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 1.
KW   Transmembrane; Leucine-rich repeat; Repeat; Glycoprotein; Signal.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29    649       LEUCINE-RICH REPEAT TRANSMEMBRANE PROTEIN
FT                                FLRT3.
FT   DOMAIN       29    528       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    529    549       POTENTIAL.
FT   DOMAIN      550    649       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       57     82       LRR 1.
FT   REPEAT       83    105       LRR 2.
FT   REPEAT      107    126       LRR 3.
FT   REPEAT      127    152       LRR 4.
FT   REPEAT      154    179       LRR 5.
FT   REPEAT      181    197       LRR 6.
FT   REPEAT      198    223       LRR 7.
FT   REPEAT      224    246       LRR 8.
```

```
FT   REPEAT      247    269       LRR 9.
FT   REPEAT      270    293       LRR 10.
FT   DOMAIN      405    485       FIBRONECTIN TYPE-III.
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    282    282       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    296    296       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    198    198       L -> P (IN REF. 3; BAB55282).
FT   CONFLICT    307    307       W -> R (IN REF. 3; BAB55282).
FT   CONFLICT    400    400       H -> Q (IN REF. 3; BAB55023).
FT   CONFLICT    519    519       K -> Q (IN REF. 3; BAB55023).
FT   CONFLICT    638    638       D -> G (IN REF. 3; BAB55282).
SQ   SEQUENCE   649 AA;  73003 MW;  9EFF666C46181F08 CRC64;

  Query Match             5.9%;  Score 240.5;  DB 1;  Length 649;
  Best Local Similarity   21.8%;  Pred. No. 3.4e-08;
  Matches  141;  Conservative   90;  Mismatches  220;  Indels  195;  Gaps   28;

Qy          3 TLLGGLLAFGMAFAVVDACPKYCVCQ----------------NLSESLGTL--------- 37
                | :|  |       :  :||  | |                  : |   ||
Db         13 TKIGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRFLTSIPTGIPEDATTLYLQNNQINN 72

Qy         38 --CPS--KGLLFV-------------PPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
                   ||  | || |             | :: :   || |  | |  |:    : :  | :
Db         73 AGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYLEE 132

Qy         81 LTLSRNTIS--HIQPFSFLDLESLR---------------------SLHLDSNRLPSLGE 117
                | |  |::|   |:  :| |   ||                      | || ||: ::
Db        133 LHLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISS 192

Qy        118 DTLRGLVNLQHLIVNNNQLG--GIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNL 175
                 :|:|| :|: |::: | |   |: |: | : |: | :| |  |:|   | :      ||
Db        193 PSLQGLTSLKRLVLDGNLLNNHGLGDKVFFN-LVNLTELSLVRNSLTAAPVNLPG--TNL 249

Qy        176 HQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPL 235
                 :| |  | :: :    |: |::| |||:::| |  | |  ||       :|
Db        250 RKLYLQDNHINRVPPNAFSYLRQLYRLDMSNNNLSNL-PQGIF--DDLDNIT-------- 298

Qy        236 SFSFGGNPLHCNCELLWLRRLERDDDLET-----------CGSPGGLKGRYFWHVREEEF 284
                      || :| |:: |:|     | |::           | :|  ::|     :  | |
Db        299 QLILRNNPWYCGCKMKWVR-----DWLQSLPVKVNVRGLMCQAPEKVRGMAIKDLNAELF 353

Qy        285 VCEP-------PLITQHTHKLLVLEGQ--AATLKCKAIGDPSPLIHWVAPDDRLVGNSSR 335
                 |:         : |   : :   :||  |   |   | :|      :  | :  |: ||
Db        354 DCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPDIKNPK-----LTKDHQTTGSPSR 408

Qy        336 TAVYDNGTLDIFITTSQ-----------------------DSGAFTCIAAN--------- 363
                  :    |: :   ||                         | ||  |
Db        409 KTI----TITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEY 464

Qy        364 --AAGEATAMVEVSIVQLP----HLSNST-----SRTAPPKSRLSDITGSSKTSRGGGGS 412
                   | |  :  :| :| :     :| : |     : ||| :      | : :  :
Db        465 LVTALEPDSPYKVCMVPMETSNLYLFDETPVCIETETAPLRMYNPTTTLNREQEK----- 519

Qy        413 GGGEPPKSP--PERAVLVSEVTTTS----ALVKWSV-------------SKSAPRVKMYQ 453
                   || |:|  |  |::   |   :    ||| | |             ||   |   |
Db        520 ---EPYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYA 576

Qy        454 LQYNCSDDEVL-----IYRMIPASN-----KAFVVNNLVSGTGYDL 489
                      |: :|      ::|:| ||     : ||:: :    | :|
Db        577 EAGTKKDNSILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNL 622


RESULT 9
LGR4_RAT
ID   LGR4_RAT       STANDARD;      PRT;   951 AA.
AC   Q9Z2H4;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
GN   GPR48 OR LGR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=99065210; PubMed=9849958;
RA   Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
RT   "Characterization of two LGR genes homologous to gonadotropin and
RT   thyrotropin receptors with extracellular leucine-rich repeats and a G
RT   protein-coupled, seven-transmembrane region.";
RL   Mol. Endocrinol. 12:1830-1845(1998).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF061443; AAC77910.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 15.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00370; LRR; 4.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 5.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Signal; Glycoprotein;
KW   Repeat; Leucine-rich repeat.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    951       LEUCINE-RICH REPEAT-CONTAINING G PROTEIN-
FT                                COUPLED RECEPTOR 4.
FT   DOMAIN       25    544       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    545    565       1 (POTENTIAL).
FT   DOMAIN      566    575       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    576    596       2 (POTENTIAL).
FT   DOMAIN      597    619       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    620    640       3 (POTENTIAL).
FT   DOMAIN      641    661       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    662    682       4 (POTENTIAL).
FT   DOMAIN      683    703       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    704    724       5 (POTENTIAL).
FT   DOMAIN      725    756       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    757    777       6 (POTENTIAL).
FT   DOMAIN      778    783       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    784    804       7 (POTENTIAL).
FT   DOMAIN      805    951       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       55     79       LRR 1.
FT   REPEAT       81    103       LRR 2.
FT   REPEAT      104    127       LRR 3.
FT   REPEAT      128    151       LRR 4.
FT   REPEAT      153    175       LRR 5.
FT   REPEAT      176    199       LRR 6.
FT   REPEAT      200    223       LRR 7.
FT   REPEAT      225    247       LRR 8.
FT   REPEAT      248    270       LRR 9.
FT   REPEAT      272    294       LRR 10.
FT   REPEAT      318    341       LRR 11.
FT   REPEAT      342    366       LRR 12.
FT   REPEAT      368    387       LRR 13.
```

```
FT   REPEAT      388    411       LRR 14.
FT   REPEAT      413    435       LRR 15.
FT   DISULFID    618    693       BY SIMILARITY.
FT   CARBOHYD     68     68       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    188    188       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    314    314       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    505    505       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   951 AA;  104138 MW;  EDD56AC072123461 CRC64;

  Query Match             5.9%;  Score 240.5;  DB 1;  Length 951;
  Best Local Similarity   22.3%;  Pred. No. 5.7e-08;
  Matches  152;  Conservative   84;  Mismatches  230;  Indels  217;  Gaps   29;

Qy          7 GLLAF---------GMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPP--------- 48
              ||| |         | : |    |   | |          |  |||  ||
Db          6 GLLCFLALGLLGSAGPSGAAPPLCAAPCSCDG---DRRVDCSGKGLTAVPEGLSAFTQAL 62

Qy         49 DIDRRTV---------------ELRLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQP 93
              ||    :               ||:| || :  |  :  : :  |  |||  |  :  :
Db         63 DISMNNITQLPEDAFKSFPFLEELQLAGNDLSLIHPKALSGLKELKVLTLQNNQLRTVPS 122

Qy         94 FSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLED 153
               :   | :|:|| ||:| : |: ||:  ||| |:|| :::| |  :      : | ||:
Db        123 EAIHGLSALQSLRLDANHITSVPEDSFEGLVQLRHLWLDDNSLTEVPVRPLSN-LPTLQA 181

Qy        154 LDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLP 213
              | |: ||:  :|  :   : :|  | | :| :  :::  |  |  |  |||  | | : |
Db        182 LTLALNNISSIPDFAFTNLSSLVVLHLHNNKIKSLSQHCFDGLDNLETLDLNYNYLDEFP 241

Qy        214 PDPIFARSQASALTATPFAPPLSF-----------SFGGNPLHCNCELLWLRRLERDDDL 262
                         |: | |    | |           :||||||        || :   |
Db        242 ----------QAIKALPSLKELGFHSNSISVIPDGAFGGNPL--------LRTIHLYD-- 281

Qy        263 ETCGSPGGLKGRYFWH---------VREEEFVCEPPLITQHTH-KLLVLEGQAATLKCKA 312
                  :|    |   :|         :|    |   | :|   | : | | |     |  :
Db        282 ----NPLSFVGNSAFHNLSDLHCLVIRGASLVQWFPNLTGTVHLESLTLTG----TKISS 333

Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
              |           ||| |  |          |||:     :|  :|            |:
Db        334 I-----------PDD-LCQNQKMLR-----TLDLSYNNIRDLPSF--------NGCRALE 368

Qy        373 EVSIV--QLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
              |:|:   |:  :  :| :      |: |:: :                        |: |
Db        369 EISLQRNQISLIKENTFQ-GLTSLRILDLSRN------------------------LIRE 403

Qy        431 V---------TTTSALVKWSVSKSAPR-----VKMYQLQYNCSDDEVLIYRMIPASNKAF 476
              :         | |:  | ::   | |      :   :|  |    : |      |:
Db        404 IHSGAFAKLGTITNLDVSFNELTSFPTEGLNGLNQLKLVGNFKLKDAL------AARDFA 457

Qy        477 VVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGC---AQFFTKADYPQCQSM-------- 525
               : :|     |  |  | |             ||   |   |: : ||  |:
Db        458 NLRSLSVPYAYQCC--AFW-------------GCDSYANLNTEDNSPQEHSVTKEKGATD 502

Qy        526 ------------HSQI-----------------LGGTMI-LVIGGIIVATLLVFIVILMV 555
                          ||||                 ||  || | :  | :  ||  :::::
Db        503 AANVTSTAENEEHSQIIIHCTPSTGAFKPCEYLLGSWMIRLTVWFIFLVALLFNLLVILT 562

Qy        556 RYKVCNHEAPSKM---AAAVSNV 575
               :  |:    ||:     :|||:
Db        563 VFASCSSLPASKLFIGLISVSNL 585

RESULT 10
FLR2_HUMAN
ID   FLR2_HUMAN     STANDARD;      PRT;   660 AA.
AC   O43155;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Leucine-rich repeat transmembrane protein FLRT2 precursor
DE   (Fibronectin-like domain-containing leucine-rich transmembrane protein
DE   2).
GN   FLRT2 OR KIAA0405.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND GYLCOSYLATION.
RX   MEDLINE=20112755; PubMed=10644439;
RA   Lacy S.E., Bonnemann C.G., Buzney E.A., Kunkel L.M.;
RT   "Identification of FLRT1, FLRT2, and FLRT3: a novel family of
RT   transmembrane leucine-rich repeat proteins.";
RL   Genomics 62:417-426(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98116655; PubMed=9455477;
RA   Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. VIII.
RT   78 new cDNA clones from brain which code for large proteins in
RT   vitro.";
RL   DNA Res. 4:307-313(1997).
CC   -!- FUNCTION: May have a function in cell adhesion and/or receptor
CC       signaling.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC   -!- TISSUE SPECIFICITY: Expressed in pancreas, skeletal muscle, brain,
CC       and heart.
CC   -!- PTM: N-glycosylated.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF169676; AAF28460.1; -.
DR   EMBL; AB007865; BAA23701.1; -.
DR   Genew; HGNC:3761; FLRT2.
DR   MIM; 604807; -.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00560; LRR; 8.
DR   Pfam; PF01462; LRRNT; 1.
DR   Pfam; PF01463; LRRCT; 1.
DR   SMART; SM00370; LRR; 3.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 1.
KW   Transmembrane; Leucine-rich repeat; Repeat; Glycoprotein; Signal.
FT   SIGNAL        1     35       POTENTIAL.
FT   CHAIN        36    660       LEUCINE-RICH REPEAT TRANSMEMBRANE PROTEIN
FT                                FLRT2.
FT   DOMAIN       36    541       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    542    562       POTENTIAL.
FT   DOMAIN      563    660       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       62     87       LRR 1.
FT   REPEAT       88    108       LRR 2.
FT   REPEAT      109    131       LRR 3.
FT   REPEAT      132    157       LRR 4.
FT   REPEAT      159    181       LRR 5.
FT   REPEAT      183    202       LRR 6.
FT   REPEAT      203    228       LRR 7.
FT   REPEAT      229    251       LRR 8.
```

```
FT   REPEAT      252    274       LRR 9.
FT   REPEAT      275    298       LRR 10.
FT   DOMAIN      420    502       FIBRONECTIN TYPE-III.
FT   CARBOHYD    202    202       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    298    298       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    433    433       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    521    521       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   660 AA;  74048 MW;  9B15F283B0D5F778 CRC64;

  Query Match             5.8%;  Score 238;  DB 1;  Length 660;
  Best Local Similarity   23.4%;  Pred. No. 5e-08;
  Matches  121;  Conservative   77;  Mismatches  206;  Indels  114;  Gaps   20;

Qy         56 ELRLGGNFIIHISRQD--FANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLP 113
              || |  | |  :  :|  |     |  | ||:| :| :     :||: ||   :| ||:
Db        137 ELHLDDNSISTVGVEDGAFREAISLKLLFLSKNHLSSVPVGLPVDLQELR---VDENRIA 193

Qy        114 SLGEDTLRGLVNLQHLIVNNNQL--GGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRR 171
               : :   : | :|: |||: | |   |||:  |   |  |::  :  |:|   | |
Db        194 VISDMAFQNLTSLERLIVDGNLLTNKGIAEGTF-SHLTKLKEFSIVRNSLSHPPPDLPG- 251

Qy        172 MVNLHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPF 231
                :| :| |  | ::||    |::|:|| |||:::|:|: |        |    |||
Db        252 -THLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTAR-- 308

Qy        232 APPLSFSFGGNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLI 291
                        ||  |:| : |:      : |:   |   ::|          |:|: |
Db        309 ---------NNPWFCDCSIKWV-----TEWLKYIPSSLNVRG----------FMCQGP-- 342

Qy        292 TQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTS 351
               :    : | |     | |                                      ||:
Db        343 -EQVRGMAVRELNMNLLSCP-------------------------------------TTT 364

Qy        352 QDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGG 411
                   ||     |   |:   :   : :|: | | :   |  |:|  |
Db        365 PGLPLFT----PAPSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIP----------- 409

Qy        412 SGGGEPPKSPP--ERAVL-VSEVTTTSALVKWSVSKSAPRVKMYQLQY----NCSDDEVL 464
               . |    :||  ||  | :  |  ||  | |    |   |  |:| :    :     ::
Db        410 DWDGRERVTPPISERIQLSIHFVNDTSIQVSW---LSLFTVMAYKLTWVKMGHSLVGGIV 466

Qy        465 IYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADY----P 520
                |::    :   : ||   : | :|:: :    |    |     |::  | | |
Db        467 QERIVSGEKQHLSLVNLEPRSTYRICLVPL---DAFNYRAVEDTICSEATTHASYLNNGS 523

Qy        521 QCQSMHSQI----LGGTMIL--VIGGIIVATLLVFIVI 552
                 | | |     :|   :|  :||| ::  |:| : :
Db        524 NTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVLLSV 561

RESULT 11
CHAD_BOVIN
ID   CHAD_BOVIN     STANDARD;      PRT;   361 AA.
AC   Q27972;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Chondroadherin precursor (Cartilage leucine-rich protein) (38 kDa bone
DE   protein).
GN   CHAD.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Cartilage;
RX   MEDLINE=94342341; PubMed=8063792;
RA   Neame P.J., Sommarin Y., Boynton R.E., Heinegaard D.;
RT   "The structure of a 38-kDa leucine-rich protein (chondroadherin)
RT   isolated from bovine cartilage.";
RL   J. Biol. Chem. 269:21547-21554(1994).
RN   [2]
RP   SEQUENCE OF 25-55 AND 77-97.
RC   TISSUE=Bone;
RX   MEDLINE=95113864; PubMed=7814406;
RA   Hu B., Coulson L., Moyer B., Price P.A.;
RT   "Isolation and molecular cloning of a novel bone phosphoprotein
RT   related in sequence to the cystatin family of thiol protease
RT   inhibitors.";
RL   J. Biol. Chem. 270:431-436(1995).
CC   -!- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and
CC       osteoblasts. This binding is mediated (at least for chondrocytes
CC       and fibroblasts) by the integrin alpha(2)beta(1). May play an
CC       important role in the regulation of chondrocyte growth and
CC       proliferation.
CC   -!- SUBUNIT: Mostly monomeric. Interacts with collagen type II (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Extracellular matrix.
CC   -!- TISSUE SPECIFICITY: Cartilage.
CC   -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCANS
CC       (SLRPS) FAMILY. CLASS IV SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U08018; AAA21330.1; -.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00560; LRR; 10.
DR   Pfam; PF01462; LRRNT; 1.
DR   Pfam; PF01463; LRRCT; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 5.
KW   Signal; Leucine-rich repeat; Repeat; Glycoprotein;
KW   Extracellular matrix.
FT   SIGNAL        1     24       OR 23 (IN SOME ISOFORM(S)).
FT   CHAIN        25    361       CHONDROADHERIN.
FT   CHAIN        25    352       CHONDROADHERIN, MINOR FORM.
FT   REPEAT       52     75       LRR 1.
FT   REPEAT       76     99       LRR 2.
FT   REPEAT      100    123       LRR 3.
FT   REPEAT      124    147       LRR 4.
FT   REPEAT      149    171       LRR 5.
FT   REPEAT      173    195       LRR 6.
FT   REPEAT      196    219       LRR 7.
FT   REPEAT      220    243       LRR 8.
FT   REPEAT      245    268       LRR 9.
FT   REPEAT      269    292       LRR 10.
FT   REPEAT      294    319       LRR 11.
FT   DISULFID     25     40       BY SIMILARITY.
FT   DISULFID    306    348
FT   DISULFID    308    328
FT   CARBOHYD    146    146       O-LINKED (POTENTIAL).
FT   CONFLICT     25     25       C -> Y (IN REF. 2).
FT   CONFLICT     29     29       C -> W (IN REF. 2).
FT   CONFLICT     31     31       C -> H (IN REF. 2).
FT   CONFLICT     40     40       C -> L (IN REF. 2).
FT   CONFLICT     52     52       S -> R (IN REF. 2).
SQ   SEQUENCE   361 AA;  40884 MW;  DA79DC98AD3DD1F8 CRC64;

  Query Match             5.8%;  Score 236;  DB 1;  Length 361;
  Best Local Similarity   25.4%;  Pred. No. 3e-08;
  Matches   90;  Conservative   42;  Mismatches  122;  Indels  100;  Gaps   11;

Qy          7 GLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIH 66
```

```
              ||||       : |||: | |   |:    :|   ||  : | :  :|  | |  |
Db         13 GLLA--SLLPALAACPQNCHCH--SDLQHVICDKVGLQKI-PKVSEKTKLLNLQRNNFPV 67

Qy         67 ISRQDFANMTGLVD------------------------LTLSRNTISHIQPFSFLDLESL 102
              ::   |  |  ||                         | || | |  ::  :| ||  |
Db         68 LATNSFRAMPNLVSLHLQHCQIREVAAGAFRGLKQLIYLYLSHNDIRVLRAGAFDDLTEL 127

Qy        103 RSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADEAFE----------------- 145
                |:|| |::  |    |  ||||  | :|||::  :   ||:
Db        128 TYLYLDHNKVTELPRGLLSPLVNLFILQLNNNKIRELRSGAFQGAKDLRWLYLSENSLSS 187

Qy        146 ------------------------------DFLLTLEDLDLSYNNLHGLPWDSVRRMVN- 174
                                              |  :|:| ||:| |  :| :: :
Db        188 LQPGALDDVENLAKFYLDRNQLSSYPSAALSKLRVVEELKLSHNPLKSIPDNAFQSFGRY 247

Qy        175 LHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPP 234
              |  | ||:  |:  ::| |  :  |  : | :||| :|| :  |       ||  |
Db        248 LETLWLDNTNLEKFSDGAFLGVTTLKHVHLENNRLHQLPSN--FPFDSLETLTLT----- 300

Qy        235 LSFSFGGNPLHCNCELLWLRR-----LERDDDLETCGSPGGLKGRYFWHVREEE 283
                     ||  | |:|  |||       | |   || ||   :|:   |:|: :
Db        301 ------NNPWKCTCQLRGLRRWLEAKTSRPD--ATCASPAKFRGQ---HIRDTD 343

RESULT 12
Y848_HUMAN
ID   Y848_HUMAN     STANDARD;      PRT;   977 AA.
AC   O94933;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein KIAA0848.
GN   KIAA0848.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=99156230; PubMed=10048485;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 5:355-364(1998).
CC   -!- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB020655; BAA74871.1; -.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00560; LRR; 9.
DR   Pfam; PF01463; LRRCT; 2.
DR   SMART; SM00370; LRR; 4.
DR   SMART; SM00082; LRRCT; 2.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 3.
KW   Hypothetical protein; Repeat; Leucine-rich repeat.
FT   REPEAT       76     99       LRR 1.
FT   REPEAT      100    123       LRR 2.
FT   REPEAT      125    147       LRR 3.
FT   REPEAT      148    171       LRR 4.
FT   REPEAT      172    195       LRR 5.
FT   REPEAT      197    219       LRR 6.
FT   REPEAT      407    430       LRR 7.
FT   REPEAT      431    454       LRR 8.
FT   REPEAT      455    478       LRR 9.
FT   REPEAT      480    502       LRR 10.
FT   REPEAT      503    526       LRR 11.
FT   REPEAT      528    550       LRR 12.
FT   DOMAIN      712    723       POLY-GLY.
SQ   SEQUENCE   977 AA;  109005 MW;  3C936B7E0003DF54 CRC64;

  Query Match             5.8%;  Score 236;  DB 1;  Length 977;
  Best Local Similarity  24.8%;  Pred. No. 1.2e-07;
  Matches  111;  Conservative   42;  Mismatches  178;  Indels  116;  Gaps  15;

Qy         21 CPKYCVCQNLSESLGTL--CPSKGLLFVPPDIDR--RTVELRLGGNFIIHISRQDFANMT 76
              ||  | |      ||    |  :|   :   : | .   :| |  | |  | | || | :
Db        373 CPTGCTCNLHINDLGLTVNCKERGFNNISELLPRPLNAKKLYLSSNLIQKIYRSDFWNFS 432

Qy         77 GLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQL 136
               |  | |  | ||::|  :|::| :|:|| |: | :  |     ||| :| :|    | :
Db        433 SLDLLHLGNNRISYVQDGAFINLPNLKSLFLNGNDIEKLTPGMFRGLQSLHYLYFEFNVI 492

Qy        137 GGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADL 196
                |   ||                           | ||  | |::|||  :    ||
Db        493 REIQPAAFS-------------------------LMPNLKLLFLNNNLLRTLPTDAFAG- 526

Qy        197 QKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELL----W 252
                ||||:|  |    ||   :     |           :      ||  | |:|:    |
Db        527 TSLARLNLRKNYFLYLPVAGVLEHLNAI----------VQIDLNENPWDCTCDLVPFKQW 576

Qy        253 LRRLER----DDDLETCGSPGGLKGRYFWHVR--EEEFVCEPPLITQHTHKLLVLEGQAA 306
              :  :       | |  | ||  | .|    ||  | | :|
Db        577 IETISSVSVVGDVL--CRSPENLTHR---DVRTIELEVLC-------------------- 611

Qy        307 TLKCKAIGDPSPLIHWVAP---------DDRLVGNSSRTAVYD----NGTLDIFITTSQD 353
                         | :  |||         |  |:|  :  : |:     | : : :
Db        612 -----------PEMLHVAPAGESPAQPGDSHLIGAPTSASPYEFSPPGGPVPLSVLILSL 660

Qy        354 SGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSR------ 407
                 |      |||    ::     :||  |         :    |:||
Db        661 LVLFFSAVFVAAGLFAYVLRRRRKKLPFRSK--------RQEGVDLTGIQMQCHRLFEDG 712

Qy        408 ---GGGGSGGGEPPKSPPERAVLVSEV 431
                 |||  ||| |  | ||:|  |  |
Db        713 GGGGGGSGGGGRPTLSSPEKAPPVGHV 739

RESULT 13
ALS_PAPHA
ID   ALS_PAPHA      STANDARD;      PRT;   605 AA.
AC   O02833;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Insulin-like growth factor binding protein complex acid labile chain
DE   precursor (ALS).
GN   IGFALS OR ALS.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=97040714; PubMed=8886027;
RA   Delhanty P., Baxter R.C.;
RT   "The cloning and expression of the baboon acid-labile subunit of the
RT   insulin-like growth factor binding protein complex.";
RL   Biochem. Biophys. Res. Commun. 227:897-902(1996).
CC   -!- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
```

```
CC       IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC   -!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
CC       IGF-I OR IGF-II AND IGFBP-3 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; S83462; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00560; LRR; 19.
DR   Pfam; PF01462; LRRNT; 1.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00370; LRR; 2.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 11.
KW   Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT   SIGNAL        1     27       BY SIMILARITY.
FT   CHAIN        28    605       INSULIN-LIKE GROWTH FACTOR BINDING
FT                                PROTEIN COMPLEX ACID LABILE CHAIN.
FT   REPEAT       53     73       LRR 1.
FT   REPEAT       74     96       LRR 2.
FT   REPEAT       98    120       LRR 3.
FT   REPEAT      121    144       LRR 4.
FT   REPEAT      146    168       LRR 5.
FT   REPEAT      169    192       LRR 6.
FT   REPEAT      194    216       LRR 7.
FT   REPEAT      217    240       LRR 8.
FT   REPEAT      242    264       LRR 9.
FT   REPEAT      265    288       LRR 10.
FT   REPEAT      289    312       LRR 11.
FT   REPEAT      313    336       LRR 12.
FT   REPEAT      338    360       LRR 13.
FT   REPEAT      361    384       LRR 14.
FT   REPEAT      386    408       LRR 15.
FT   REPEAT      409    432       LRR 16.
FT   REPEAT      433    456       LRR 17.
FT   REPEAT      458    480       LRR 18.
FT   REPEAT      482    504       LRR 19.
FT   REPEAT      505    530       LRR 20.
FT   CARBOHYD     64     64       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     85     85       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     96     96       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    515    515       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    580    580       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   605 AA;  66110 MW;  9D71994625F23652 CRC64;

  Query Match             5.8%;  Score 235.5;  DB 1;  Length 605;
  Best Local Similarity   30.5%;  Pred. No. 6.5e-08;
  Matches   78;  Conservative   34;  Mismatches  111;  Indels   33;  Gaps    4;

Qy         20 ACPKYCVCQNLSE--SLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTG 77
              |||  | |    |   |   | |: |  :|  |   |  | |  | :  |    | |::
Db         40 ACPATCACSYDDEVNELSVFCSSRNLTRLPDGIPGGTQALWLDSNNLSSIPPAAFRNLSS 99

Qy         78 LVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLG 137
              |  | |    :  ::| : | ||:|  |||: |:| ||   |      |  | ::||:|
Db        100 LAFLNLQGGQLGSLEPQALLGLENLCHLHLERNQLRSLAVGTFAYTPALALLGLSNNRLS 159

Qy        138 GIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQ 197
               : |  ||  |  | ||:| :|:|  ||  : | :  | :| |  | | ::    |: |
Db        160 RLEDGLFEG-LGNLWDLNLGWNSLAVLPDAAFRGLGGLRELVLAGNRLAYLQPALFSGLA 218

Qy        198 KLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLE 257
              :|  |||: | |:              |: |  ||                :|  |::|
Db        219 ELRELDLSRNALR--------------AIKANVFA----------------QLPRLQKLY 248

Qy        258 RDDDLETCGSPGGLKG 273
               | :|    :||   |
Db        249 LDRNLIAAVAPGAFLG 264


RESULT 14
LGR4_HUMAN
ID   LGR4_HUMAN     STANDARD;      PRT;   951 AA.
AC   Q9BXB1; Q9NYD1;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
DE   (G protein-coupled receptor 48).
GN   GPR48 OR LGR4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=21294803; PubMed=11401528;
RA   Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
RT   "Molecular characterization of a novel glycoprotein hormone
RT   G-protein-coupled receptor.";
RL   Biochem. Biophys. Res. Commun. 282:757-764(2001).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
CC       placenta, ovary, testis and adrenal. Expressed also in spinal
CC       cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate
CC       and spleen.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF346711; AAK31153.1; -.
DR   EMBL; AF346709; AAK31153.1; JOINED.
DR   EMBL; AF346710; AAK31153.1; JOINED.
DR   EMBL; AF257182; AAF68989.1; -.
DR   Genew; HGNC:13299; GPR48.
DR   MIM; 606666; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 15.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00370; LRR; 6.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 15.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Signal; Transmembrane; Glycoprotein;
```

```
KW   Repeat; Leucine-rich repeat.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    951       LEUCINE-RICH REPEAT-CONTAINING G PROTEIN-
FT                                COUPLED RECEPTOR 4.
FT   DOMAIN       25    544       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    545    565       1 (POTENTIAL).
FT   DOMAIN      566    575       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    576    596       2 (POTENTIAL).
FT   DOMAIN      597    620       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    621    641       3 (POTENTIAL).
FT   DOMAIN      642    661       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    662    682       4 (POTENTIAL).
FT   DOMAIN      683    703       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    704    724       5 (POTENTIAL).
FT   DOMAIN      725    756       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    757    777       6 (POTENTIAL).
FT   DOMAIN      778    783       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    784    804       7 (POTENTIAL).
FT   DOMAIN      805    951       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       55     79       LRR 1.
FT   REPEAT       81    103       LRR 2.
FT   REPEAT      104    127       LRR 3.
FT   REPEAT      128    151       LRR 4.
FT   REPEAT      153    175       LRR 5.
FT   REPEAT      176    199       LRR 6.
FT   REPEAT      201    223       LRR 7.
FT   REPEAT      224    247       LRR 8.
FT   REPEAT      248    270       LRR 9.
FT   REPEAT      272    294       LRR 10.
FT   REPEAT      318    341       LRR 11.
FT   REPEAT      342    366       LRR 12.
FT   REPEAT      368    387       LRR 13.
FT   REPEAT      388    411       LRR 14.
FT   REPEAT      413    435       LRR 15.
FT   DISULFID    618    693       BY SIMILARITY.
FT   CARBOHYD     68     68       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    314    314       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    505    505       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    292    292       F -> S (IN REF. 1; AAF68989).
FT   CONFLICT    433    433       L -> P (IN REF. 1; AAF68989).
FT   CONFLICT    668    668       L -> S (IN REF. 1; AAF68989).
SQ   SEQUENCE   951 AA;  104460 MW;  5E0C2DFCF22CA1BB CRC64;

  Query Match             5.7%;  Score 234.5;  DB 1;  Length 951;
  Best Local Similarity   22.0%;  Pred. No. 1.4e-07;
  Matches  148;  Conservative   79;  Mismatches  251;  Indels  195;  Gaps   27;

Qy          7 GLLAF---------GMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVEL 57
              ||| |         | : |    |   | |          |  |||  ||  :   |  |
Db          6 GLLCFLALGLLGSAGPSGAAPPLCAAPCSCDG---DRRVDCSGKGLTAVPEGLSAFTQAL 62

Qy         58 RLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLE----------------- 100
               :  | |  :    | |   | :| |: | :| | | :   |:
Db         63 DISMNNITQLPEDAFKNFPFLEELQLAGNDLSFIHPKALSGLKELKVLTLQNNQLKTVPS 122

Qy        101 -------SLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLED 153
                     :|:|| ||:| : |: ||:  ||| |:|| :::| |  :      : | ||:
Db        123 EAIRGLSALQSLRLDANHITSVPEDSFEGLVQLRHLWLDDNSLTEVPVHPLSN-LPTLQA 181

Qy        154 LDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLP 213
              | |: | :  :|  :   : :|  | | :| :  :::  |  |  |  |||: | | : |
Db        182 LTLALNKISSIPDFAFTNLSSLVVLHLHNNKIRGLSQHCFDGLDNLETLDLSYNNLGEFP 241

Qy        214 PDPIFARSQASALTATPFAPPLSF-----------SFGGNPLHCNCELLWLRRLERDDDL 262
                         |: | |    | |           :| ||||        || :   |
Db        242 ----------QAIKARPSLKELGFHSNSISVIPDGAFDGNPL--------LRTIHLYD-- 281

Qy        263 ETCGSPGGLKGRYFWH----------VREEEFVCEPPLITQHTH-KLLVLEGQAATLKCKA 312
                  :|    |   :|         :|    | : | :|   | : | | |       |
Db        282 ----NPLSFVGNSAFHNLSDLHSLVIRGASMVQQFPNLTGTVHLESLTLTG------TKI 331
```

```
Qy        313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFT-CIAAN-------- 363
                 |: |      : :::            |||:     :|  :|  | |
Db        332 SSIPNNL----CQEQKML-----------RTLDLSYNNIRDLPSFNGCHALEEISLQRNQ 376

Qy        364 ----AAGEATAMVEVSIVQLP-----HLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGG 414
                    |    :: : |: |       : :    |  | : | |:: :  ||    |  |
Db        377 IYQIKEGTFQGLISLRILDLSRNLIHEIHSRAFATLGPITNL-DVSFNELTSFPTEGLNG 435

Qy        415 GEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNK 474
                  |       ||       ||     :|    ::   : |
Db        436 LNQLK-------LVGNFKLKEAL----AAKDFVNLRSLSVPY------------------ 466

Qy        475 AFVVNNLVSGTGYDLCVLAMW--DDTATTLTATNIV---GCAQFFTKADYPQCQSM---- 525
                          |  |  | |  |  |   |  | :     ||    ||     |
Db        467 -----------AYQCC--AFWGCDSYANLNTEDNSLQDHSVAQEKGTADAANVTSTLENE 513

Qy        526 -HSQI-----------------LGGTMI-LVIGGIIVATLLVFIVILMVRYKVCNHEAPS 566
               ||||                 ||  || | :  | :  |   :::::  :  |     |
Db        514 EHSQIIIHCTPSTGAFKPCEYLLGSWMIRLTVWFIFLVALFFNLLVILTTFASCTSLPSS 573

Qy        567 KM---AAAVSNVY 576
              |:     :|||::
Db        574 KLFIGLISVSNLF 586


RESULT 15
LGR7_HUMAN
ID   LGR7_HUMAN     STANDARD;      PRT;   757 AA.
AC   Q9HBX9;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Relaxin receptor 1 (Leucine-rich repeat-containing G protein-coupled
DE   receptor 7).
GN   LGR7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF ASP-637.
RX   MEDLINE=20388592; PubMed=10935549;
RA   Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA   van der Spek P.J., van Duin M., Hsueh A.J.W.;
RT   "The three subfamilies of leucine-rich repeat-containing G protein-
RT   coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT   signaling mechanism for LGR7.";
RL   Mol. Endocrinol. 14:1257-1271(2000).
RN   [2]
RP   CHARACTERIZATION.
RX   PubMed=11517286;
RA   Bartsch O., Bartlick B., Ivell R.;
RT   "Relaxin signalling links tyrosine phosphorylation to
RT   phosphodiesterase and adenylyl cyclase activity.";
RL   Mol. Hum. Reprod. 7:799-809(2001).
CC   -!- FUNCTION: Receptor for relaxin. The activity of this receptor is
CC       mediated by G proteins leading to stimulation of adenylate cyclase
CC       and an increase of cAMP. Binding of the ligand may also
CC       activate a tyrosine kinase pathway that inhibits the activity of a
CC       phosphodiesterase that degrades cAMP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: Expressed in the brain, kidney, testis,
CC       placenta, uterus, ovary, adrenal, prostate, skin and heart. Not
CC       detected in spleen.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC   -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
```

```
CC   use  by  non-profit  institutions as long  as its content  is  in  no 'way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF190500; AAG17167.1; -.
DR   MIM; 606654; -.
DR   HSSP; P01130; 1LDR.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_recept_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 1.
DR   Pfam; PF00560; LRR; 8.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00192; LDLa; 1.
DR   SMART; SM00370; LRR; 7.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 10.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat; Alternative splicing.
FT   DOMAIN        1    409       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    410    430       1 (POTENTIAL).
FT   DOMAIN      431    443       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    444    464       2 (POTENTIAL).
FT   DOMAIN      465    486       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    487    507       3 (POTENTIAL).
FT   DOMAIN      508    527       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    528    548       4 (POTENTIAL).
FT   DOMAIN      549    577       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    578    598       5 (POTENTIAL).
FT   DOMAIN      599    629       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    630    650       6 (POTENTIAL).
FT   DOMAIN      651    651       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    652    672       7 (POTENTIAL).
FT   DOMAIN      673    757       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       26     63       LDL-RECEPTOR CLASS A.
FT   REPEAT      125    148       LRR 1.
FT   REPEAT      150    172       LRR 2.
FT   REPEAT      173    196       LRR 3.
FT   REPEAT      198    220       LRR 4.
FT   REPEAT      221    244       LRR 5.
FT   REPEAT      246    269       LRR 6.
FT   REPEAT      270    293       LRR 7.
FT   REPEAT      294    317       LRR 8.
FT   REPEAT      319    341       LRR 9.
FT   REPEAT      343    365       LRR 10.
FT   DISULFID    485    563       BY SIMILARITY.
FT   CARBOHYD     36     36       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    264    264       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    272    272       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     63     96       GDNNGWSMQFDKYFASYYKMTSQYPFEAETPECL -> V
FT                                (IN ISOFORM 2).
FT   MUTAGEN     637    637       D->Y: LEADS TO CONSTITUTIVE INCREASE OF
FT                                BASAL CAMP.
SQ   SEQUENCE   757 AA;  86992 MW;  8079E8DBF3A3EF21 CRC64;

  Query Match             5.7%;  Score 232.5;  DB 1;  Length 757;
  Best Local Similarity   31.8%;  Pred. No. 1.4e-07;
  Matches   69;  Conservative   34;  Mismatches   95;  Indels   19;  Gaps    3;

Qy         17 VVDACPKYCVCQNLS-----------ESLGTLCPSKGLLF-----VPPDIDRRTVELR-- 58
              :| : |  |:|| |             |: :   :  | :     :|||  :   :|:
```

```
Db         96 LVGSVPVQCLCQGLELDCDETNLRAVPSVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKL 155

Qy         59 -LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGE 117
               |  | |  ||   |  :  |  | || | |: ::|  | ||  |  | :: | |  :
Db        156 YLQNNKITSISIYAFRGLNSLTKLYLSHNRITFLKPGVFEDLHRLEWLIIEDNHLSRISP 215

Qy        118 DTLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQ 177
               |  || :|  |:: || |  : |:     :  |  |||  |::| |   :     ||
Db        216 PTFYGLNSLILLVLMNNVLTRLPDKPLCQHMPRLHWLDLEGNHIHNLRNLTFISCSNLTV 275

Qy        178 LSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPP 214
              | :  | ::|: | ||| ||||  ||| ||::: |||
Db        276 LVMRKNKINHLNENTFAPLQKLDELDLGSNKIENLPP 312

Search completed: June  8, 2003, 21:29:07
Job time : 41 secs
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         June  8, 2003, 20:50:18 ; Search time 62 Seconds
                                          (without alignments)
                                          2622.116 Million cell updates/sec

Title:          US-09-831-846-2
Perfect score:  4094
Sequence:       1 METLLGGLLAFGMAFAVVDA..........DLVGARGTFGSSEWVMESTV 789

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:      671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_21:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                        Description
----------------------------------------------------------------------------
    1     4094   100.0    832  4  Q9ULH4                    Q9ulh4 homo sapien
    2     4075    99.5    789  6  Q9BE71                    Q9be71 macaca fasc
    3   3885.5    94.9    788  11  Q9CYK3                    Q9cyk3 mus musculu
    4     1821    44.5    719  4  Q96NI6                    Q96ni6 homo sapien
    5     1573    38.4    628  4  Q9BTN0                    Q9btn0 homo sapien
    6   1571.5    38.4    700  4  Q9P244                    Q9p244 homo sapien
    7   1320.5    32.3    492  11  Q99KT6                    Q99kt6 mus musculu
    8   1063.5    26.0    542  4  Q8TBS9                    Q8tbs9 homo sapien
    9      775    18.9    450  4  Q96C50                    Q96c50 homo sapien
   10      653    16.0    324  4  Q9BWJ0                    Q9bwj0 homo sapien
   11    435.5    10.6    640  4  Q9HCJ2                    Q9hcj2 homo sapien
   12    418.5    10.2    606  4  Q9BZ20                    Q9bz20 homo sapien
   13      385     9.4    649  4  Q96A85                    Q96a85 homo sapien
   14    383.5     9.4    648  11  Q8VI35                    Q8vi35 mus musculu
   15      383     9.4    653  4  Q9HBW1                    Q9hbw1 homo sapien
   16    380.5     9.3    707  11  Q9ESY6                    Q9esy6 rattus norv
```

```
   17      380     9.3    614  6  Q9N008                    Q9n008 macaca fasc
   18      379     9.3    614  4  Q96FE5                    Q96fe5 homo sapien
   19      374     9.1    614  11  Q9D1T0                    Q9d1t0 mus musculu
   20      373     9.1    707  11  P97860                    P97860 mus musculu
   21      368     9.0    708  4  Q9H3W5                    Q9h3w5 homo sapien
   22      368     9.0   1091  11  P70193                    P70193 mus musculu
   23      365     8.9   1496  4  Q92626                    Q92626 homo sapien
   24    362.5     8.9   1093  4  Q96JA1                    Q96ja1 homo sapien
   25    361.5     8.8   1094  4  Q9BYB8                    Q9byb8 homo sapien
   26    354.5     8.7    718  13  O73675                    O73675 xenopus lae
   27    350.5     8.6    422  4  Q9NT99                    Q9nt99 homo sapien
   28    341.5     8.3    892  5  Q9VWI6                    Q9vwi6 drosophila
   29    334.5     8.2    730  4  Q9P231                    Q9p231 homo sapien
   30    331.5     8.1    716  11  Q61809                    Q61809 mus musculu
   31    322.5     7.9   1065  4  O94898                    O94898 homo sapien
   32      312     7.6   1021  5  Q9V430                    Q9v430 drosophila
   33      304     7.4    473  6  Q9N0E3                    Q9n0e3 macaca fasc
   34      303     7.4    473  4  Q9BZR6                    Q9bzr6 homo sapien
   35      296     7.2    441  11  Q99PH1                    Q99ph1 mus musculu
   36      295     7.2    473  11  Q99PI8                    Q99pi8 mus musculu
   37      295     7.2   1316  5  Q9VQ25                    Q9vq25 drosophila
   38      294     7.2    836  5  Q9V9V6                    Q9v9v6 drosophila
   39    292.5     7.1    892  5  P91644                    P91644 drosophila
   40    290.5     7.1    894  5  Q9VKG1                    Q9vkg1 drosophila
   41    290.5     7.1   1530  11  Q9WUG5                    Q9wug5 rattus norv
   42    286.5     7.0   1531  11  O88279                    O88279 rattus norv
   43    284.5     6.9   1173  5  Q9V7J8                    Q9v7j8 drosophila
   44    284.5     6.9   1525  4  Q9Y5Q7                    Q9y5q7 homo sapien
   45    283.5     6.9    798  4  Q8WWZ2                    .Q8wwz2 homo sapien

                                  ALIGNMENTS

RESULT 1
Q9ULH4
ID   Q9ULH4     PRELIMINARY;      PRT;   832 AA.
AC   Q9ULH4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   KIAA1246 protein (Fragment).
GN   KIAA1246.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE=20039619; PubMed=10574462;
RA   Nagase T., Ishikawa K., Kikuno R., Hirosawa M., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
DR   EMBL; AB033072; BAA86560.1; -.
DR   HSSP; P56276; 1TLK.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 7.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00060; FN3; 1.
DR   SMART; SM00408; IGc2; 1.
```

```
DR   SMART; SM00370; LRR; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 3.
KW   Immunoglobulin domain.
FT   NON_TER       1      1
SQ   SEQUENCE    832 AA;  88978 MW;  9B0E80F7BFF974A8 CRC64;

  Query Match             100.0%;  Score 4094;  DB 4;  Length 832;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-310;
  Matches  789;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         44 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 103

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 163

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 223

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 283

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 343

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 403

Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 463

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 523

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 583

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 643

Qy        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 703

Qy        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 763

Qy        721 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 823

Qy        781 SEWVMESTV 789
              |||||||||
Db        824 SEWVMESTV 832


RESULT 2
Q9BE71
ID   Q9BE71      PRELIMINARY;      PRT;   789 AA.
AC   Q9BE71;
```

```
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 84.7 kDa protein.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=FRONTAL LOBE LEFT;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB056799; BAB39323.1; -.
DR   HSSP; P56276; 1TLK.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 7.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00060; FN3; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00370; LRR; 5.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 6.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   789 AA;  84730 MW;  BB86DE81BC284B23 CRC64;

  Query Match              99.5%;  Score 4075;  DB 6;  Length 789;
  Best Local Similarity    99.4%;  Pred. No. 1.5e-308;
  Matches  784;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360

Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 AANAAGEATATVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
              ||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||
Db        421 PPERAVLVSEVTTTSALAKWSVSKSTPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600

Qy        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
              |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        601 KVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRLPPSAPRPKPSLDRLMGAF 660

Qy        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKR 720

Qy        721 SHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780
              ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db        721 SHSFDMGDFAAAAAGGVVPGGYSPPRRVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780

Qy        781 SEWVMESTV 789
              |||||||||
Db        781 SEWVMESTV 789


RESULT 3
Q9CYK3
ID   Q9CYK3      PRELIMINARY;      PRT;   788 AA.
AC   Q9CYK3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   5730420O05Rik protein.
GN   5730420O05RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK017594; BAB30828.1; -.
DR   HSSP; P56276; 1TLK.
DR   MGD; MGI:1917780; 5730420O05Rik.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 7.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00060; FN3; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 6.
KW   Immunoglobulin domain.
SQ   SEQUENCE   788 AA;  85076 MW;  70B70C5E6B52277B CRC64;

  Query Match             94.9%;  Score 3885.5;  DB 11;  Length 788;
  Best Local Similarity   94.9%;  Pred. No. 8.6e-294;
  Matches  750;  Conservative   10;  Mismatches   27;  Indels    3;  Gaps    2;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKRLLFVPPDIDRRTVELRLG 60

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNFIIHIGRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        121 RGLVNLQHLIVNNNQLGGIADDAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASLLTATPFAPPLSFSFG 240

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              ||||||||||||||||||||||||||||| ||||||||:|||||||||||||||||||||
Db        241 GNPLHCNCELLWLRRLERDDDLETCGSPGSLKGRYFWHIREEEFVCEPPLITQHTHKLLV 300

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||
Db        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDILITTSQDSGPFTCI 360

Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              |||||||||| ||||||||||||||||| |||||||||||||||||||||||| ||||||
Db        361 AANAAGEATATVEVSIVQLPHLSNSTSRMAPPKSRLSDITGSSKTSRGGGGSGAGEPPKS 420

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
               ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TPERAVLVSDVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQIKGGTMILVIGG 540

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAA-VSNVYSQTNGAQPPPPSSAPAGAPPQGP 599
              ||||||||||||||||||||||: | ||||| ||||||||||:||||    | |  || |
Db        541 IIVATLLVFIVILMVRYKVCNHDTPGKMAAATVSNVYSQTNGSQPPPLGGIPVGQLPQAP 600

Qy        600 PKVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGA 659
              |||||||||:||: ||||| |||||||||||||||||| |||:|| ||||||||||||||
Db        601 PKVVVRNELMDFSTSLARACDSSSSSSLGSGEAAGLGRGPWRLPPPAPRPKPSLDRLMGA 660

Qy        660 FASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAK 719
              ||||||||||||||||||||||||||||:||||||||||||||| |||||||||||||||
Db        661 FASLDLKSQRKEELLDSRTPAGRGAGTSSRGHHSDREPLLGPPATRARSLLPLPLEGKAK 720

Qy        720 RSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFG 779
```

```
              |||||||||||||||   |||||||||:||||||||||||||||||||||||||||||||
Db        721 RSHSFDMGDFAAAAA--AVPGGYSPPRRVSNIWTKRSLSVNGMLLPFEESDLVGARGTFG 778

Qy        780 SSEWVMESTV 789
              ||||||||||
Db        779 SSEWVMESTV 788


RESULT 4
Q96NI6
ID   Q96NI6       PRELIMINARY;      PRT;   719 AA.
AC   Q96NI6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CDNA FLJ30803 fis, clone FEBRA2001245, weakly similar to NAG14.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA   Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA   Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA   Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA   Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK055365; BAB70910.1; -.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR001005; Myb_DNA_binding.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 6.
DR   Pfam; PF01463; LRRCT; 1.
DR   PROSITE; PS00037; MYB_1; UNKNOWN_1.
SQ   SEQUENCE   719 AA;  79472 MW;  CE301F219705FF4B CRC64;

  Query Match             44.5%;  Score 1821;  DB 4;  Length 719;
  Best Local Similarity   51.4%;  Pred. No. 4.6e-133;
  Matches  375;  Conservative  118;  Mismatches  195;  Indels   42;  Gaps   13;

Qy          1 METLLGGLLAFGMAFAVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLG 60
              || :|  |   |:|      ||| |||| || :| |||  ||||||||:||||||||||
Db          1 MEKILFYLFLIGIA-VKAQICPKRCVCQILSPNLATLCAKKGLLFVPPNIDRRTVELRLA 59

Qy         61 GNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTL 120
               ||: :| |:|||||| |||||||||||| | | :| || :||:|||:||||  :  |
Db         60 DNFVTNIKRKDFANMTSLVDLTLSRNTISFITPHAFADLRNLRALHLNSNRLTKITNDMF 119

Qy        121 RGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSL 180
               || || |||:|||||  |:  ||:| :  ||:||||||||  :|||:| :||:|| |||
Db        120 SGLSNLHHLILNNNQLTLISSTAFDD-VFALEELDLSYNNLETIPWDAVEKMVSLHTLSL 178

Qy        181 DHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFG 240
              |||::|:| :|||: | |: |||:|||:||||||||:| |:|  | :        : |||
Db        179 DHNMIDNIPKGTFSHLHKMTRLDVTSNKLQKLPPDPLFQRAQVLATSGIISPSTFALSFG 238

Qy        241 GNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLV 300
              |||||||||||||||| |:|||||| ||  | ||||| : ||||:|||||||:|||:: |
Db        239 GNPLHCNCELLWLRRLSREDDLETCASPPLLTGRYFWSIPEEEFLCEPPLITRHTHEMRV 298

Qy        301 LEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
              |||| |||:||| ||| | |||::|: :|: |::|: ||||||||| ||| :|:||||||
Db        299 LEGQRATLRCKARGDPEPAIHWISPEGKLISNATRSLVYDNGTLDILITTVKDTGAFTCI 358
```

```
Qy        361 AANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKS 420
              |:| ||||| :|:: |::|||| |||:    |    |||: |:|:      | |
Db        359 ASNPAGEATQIVDLHIIKLPHLLNSTNHIHEPDPGSSDISTSTKSGSNTSSSNG---DTK 415

Qy        421 PPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNN 480
                :  ::|:| |:::||:|::  :: | ::|:|:||| : |: |:||||| ::| |:|||
Db        416 LSQDKIVVAEATSSTALLKFNFQRNIPGIRMFQIQYNGTYDDTLVYRMIPPTSKTFLVNN 475

Qy        481 LVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGG 540
              | :|| |||||||::||  |:|||| :||| || |: || :|  | || ||||||::|||
Db        476 LAAGTMYDLCVLAIYDDGITSLTATRVVGCIQFTTEQDYVRCHFMQSQFLGGTMIIIIGG 535

Qy        541 IIVATLLVFIVILMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPP 600
              ||||::||||:|||:||||||:    |:   ||||||||||||        :   ||
Db        536 IIVASVLVFIIILMIRYKVCNNNGQHKV-TKVSNVYSQTNGAQ----IQGCSVTLPQSVS 590

Qy        601 KVVVRNELLDFTASLARASDSS---SSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLM 657
              |  | :|     |   :|:  :   || :  | :::    |   :|||
Db        591 KQAVGHE---ENAQCCKATSDNVIQSSETCSSQDSSTTTSA---LPPS-----------W 633

Qy        658 GAFASLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLP---- 713
               :  |:  | :||     |  |    | |:    :::|        : |  |:  |
Db        634 TSSTSVSQKQKRKTGTKPSTEPQNE-AVTNVESQNTNRN------NSTALQLVSRPPDSV 686

Qy        714 LEG-KAKRSH 722
               ||  :||:|
Db        687 TEGPTSKRAH 696


RESULT 5
Q9BTN0
ID   Q9BTN0       PRELIMINARY;      PRT;   628 AA.
AC   Q9BTN0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 66.3 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC003578; AAH03578.1; -.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 6.
DR   Pfam; PF01463; LRRCT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00060; FN3; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 6.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   628 AA;  66259 MW;  FEDC5A5056ABC5FC CRC64;

  Query Match             38.4%;  Score 1573;  DB 4;  Length 628;
  Best Local Similarity   53.1%;  Pred. No. 7.8e-114;
  Matches  312;  Conservative   88;  Mismatches  174;  Indels   14;  Gaps    5;
```

```
Qy           16 AVVDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANM 75
                |    ||: | ||  |  |  |||  ||||||| :|||  ||||  |||  : |:| |||
Db           23 ATPSPCPRRCRCQTQSLPLSVLCPGAGLLFVPPSLDRRAAELRLADNFIASVRRRDLANM 82

Qy           76 TGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQ 135
                |||: |:|||||| |:   :| || :||:|||| ||| ||||  |||||||:|||::|||
Db           83 TGLLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLILSNNQ 142

Qy          136 LGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFAD 195
                |  :|  | :|   ||||||||||||  |||::: |: |:: | ||||||  :  | |:
Db          143 LAALAAGALDDCAETLEDLDLSYNNLEQLPWEALGRLGNVNTLGLDHNLLASVPAGAFSR 202

Qy          196 LQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLS---FSFGGNPLHCNCELLW 252
                | ||||||:|||||  :||||:|:|     | | |   | |    :||||||||||||:|
Db          203 LHKLARLDMTSNRLTTIPPDPLFSR---LPLLARPRGSPASALVLAFGGNPLHCNCELVW 259

Qy          253 LRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKA 312
                |||| |:|||| | ||  | ||||| | |||||||||::|  :  | |  |: | |:|:|
Db          260 LRRLAREDDLEACASPPALGGRYFWAVGEEEFVCEPPVVTHRSPPLAVPAGRPAALRCRA 319

Qy          313 IGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
                :||| | : ||:|  ||:|||||   : ||||:: :|   | | |||||||||||||| |
Db          320 VGDPEPRVRWVSPQGRLLGNSSRARAFPNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 379

Qy          373 EVSI--VQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSE 430
                |:::     | |:||||   |       :|  |  |     :  |     | :| | |:|
Db          380 ELTVGPPPPPQLANSTSCDPPRDGDPDALTPPSAASASAKVADTG----PPTDRGVQVTE 435

Qy          431 VTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLC 490
                   |:|||:|   :  | ::|||:||| | |::|:|||||| :::|:: :| ||  ||||
Db          436 HGATAALVQWPDQRPIPGIRMYQIQYNSSADDILVYRMIPAESRSFLLTDLASGRTYDLC 495

Qy          491 VLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFI 550
                |||:::|:|| ||||  ||||:| |:     | : |:  ||||||: :||:|||::||||
Db          496 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGGTMIIALGGVIVASVLVFI 555

Qy          551 VILMVRYKVCNHEAP--SKMAAAVSNVYSQTNGAQPPPPSSAPAGAPP 596
                 :|::||||   : |  :|: | ||:| ||||||  | |: ||    |
Db          556 FVLLMRYKVHGGQPPGKAKIPAPVSSVCSQTNGALGPTPTPAPPAPEP 603

RESULT 6
Q9P244
ID   Q9P244      PRELIMINARY;      PRT;   700 AA.
AC   Q9P244;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   KIAA1484 protein (Fragment).
GN   KIAA1484.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20277482; PubMed=10819331;
RA   Nagase T., Kikuno R., Ishikawa K., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human
RT   genes.XVII.The complete sequences of 100 new cDNA clones from brain
RT   which code for large proteins in vitro.";
RL   DNA Res. 7:143-150(2000).
DR   EMBL; AB040917; BAA96008.1; -.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   InterPro; IPR001005; Myb_DNA_binding.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 6.
DR   Pfam; PF01463; LRRCT; 1.
DR   SMART; SM00060; FN3; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00370; LRR; 4.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00369; LRR_TYP; 1.
DR   PROSITE; PS00037; MYB_1; UNKNOWN_1.
KW   Immunoglobulin domain.
FT   NON_TER       1      1
SQ   SEQUENCE   700 AA;  74865 MW;  30D0432242494159 CRC64;

  Query Match             38.4%;  Score 1571.5;  DB 4;  Length 700;
  Best Local Similarity   45.7%;  Pred. No. 1.2e-113;
  Matches  352;  Conservative  106;  Mismatches  203;  Indels  109;  Gaps   18;

Qy           59 LGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGED 118
                |  |||  : |:|||||| || ||||||||  :   :| || :||:||||||||  :  |
Db            1 LTDNFIAAVRRRDFANMTSLVHLTLSRNTIGQVAAGAFADLRALRALHLDSNRLAEVRGD 60

Qy          119 TLRGLVNLQHLIVNNNQLGGIADEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQL 178
                 |||| ||:|||: |||:  :   ||: || |:||||||||||  |||::| :||||: |
Db           61 QLRGLGNLRHLILGNNQIRRVESAAFDAFLSTVEDLDLSYNNLEALPWEAVGQMVNLNTL 120

Qy          179 SLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAP-PLSF 237
                :|||||:||||||||  | || |||:||||| ||||| :| ||| :     |  | ||:
Db          121 TLDHNLIDHIAEGTFVQLHKLVRLDMTSNRLHKLPPDGLFLRSQGTG----PKPPTPLTV 176

Qy          238 SFGGNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRYFWHVREEEFVCEPPLIT-QHTH 296
                ||||||||||||||||||| |:||||||  :|  |  |||| : ||||:||||||| |
Db          177 SFGGNPLHCNCELLWLRRLTREDDLETCATPEHLTDRYFWSIPEEEFLCEPPLITRQAGG 236

Qy          297 KLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGA 356
                : ||:|||| :|:|:|:||| |::|||||| ||:|||||| |  :||||: ||| :|||
Db          237 RALVVEGQAVSLRCRAVGDPEPVVHWVAPDGRLLGNSSRTRVRGDGTLDVTITTLRDSGT 296

Qy          357 FTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGE 416
                |||||:|||||||| ||| :| || :       ||| :    :|    :     |  |
Db          297 FTCIASNAAGEATAPVEVCVVPLPLM-------APPPAAPPPLTEPGSSDIATPGRPGAN 349

Qy          417 PPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAF 476
                   |  || :: :|:|: | |::|   :  | ::|||:||| | |: |:|||||:::: |
Db          350 --DSAAERRLVAAELTSNSVLIRWPAQRPVPGIRMYQVQYNSSVDDSLVYRMIPSTSQTF 407

Qy          477 VVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMIL 536
                :||:| :|  |||||||::|| || | || :|||.|| |  |   |: : :  ||||||:
Db          408 LVNDLAAGRAYDLCVLAVYDDGATALPATRVVGCVQFTTAGDPAPCRPLRAHFLGGTMII 467

Qy          537 VIGGIIVATLLVFIVILMVRYKV-----CNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAP 591
                 |||:|||::|||||:||:||||           |:    ||:| ||||||      :|
Db          468 AIGGVIVASVLVFIVLLMIRYKVYGDGDSRRVKGSRSLPRVSHVCSQTNGA---GTGAAQ 524

Qy          592 AGAPP-------------QGPPKVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRA 638
                | | |             |  | | |           |:| :: ::|:
Db          525 APALPAQDHYEALREVESQAAPAVAVE----------AKAMEAETASA------------ 562

Qy          639 PWRIPPSAPRPKPSLDRLMGAFA-SLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREP 697
                          |:  | | :|  | || |    :    :||         :| |    |
Db          563 ---------EPEVVLGRSLGGSATSLCLLPSEETSGEESR---------AAVGPRRSRSG 604

Qy          698 LLGPPAARARSLLPLP----LEGKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWT 753
                 | || :   :|  :|       : :: :||| ||:      | :   :| ||:     |
Db          605 ALEPPTSAPPTLALVPGGAAARPRPQQRYSFD-GDY------GALFQSHSYPRRARR--T 655

Qy          754 KRSLSVNGMLLPFEESDLVGARG--------------TFGSSEWVMESTV 789
                ||  |      |  :    || |               | |:||::||||
Db          656 KRHRST-----PHLDGAGGGAAGEDGDLGLGSARACLAFTSTEWMLESTV 700

RESULT 7
Q99KT6
ID   Q99KT6      PRELIMINARY;      PRT;   492 AA.
```

```
AC   Q99KT6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 53.3 kDa protein.
GN   BC004018.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC004018; AAH04018.1; -.
DR   MGD; MGI:2136810; BC004018.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   InterPro; IPR001005; Myb_DNA_binding.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 6.
DR   Pfam; PF01463; LRRCT; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 6.
DR   PROSITE; PS00037; MYB_1; UNKNOWN_1.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   492 AA;  53341 MW;  0EB01D3772B24327 CRC64;

  Query Match             32.3%;  Score 1320.5;  DB 11;  Length 492;
  Best Local Similarity   57.6%;  Pred. No. 2.5e-94;
  Matches  258;  Conservative   62;  Mismatches  113;  Indels   15;  Gaps    5;

Qy           21 CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
                ||  |:|||:: :|  ||   ||||||| |||| |||||  |||  : |:|||||| ||
Db           34 CPGRCICQNVAPTLTMLCAKTGLLFVPPAIDRRVVELRLTDNFIAAVRRRDFANMTSLVH 93

Qy           81 LTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIA 140
                ||||||||  :   :| || :||:||||||||  :  | |||| ||:|||: |||:  :
Db           94 LTLSRNTIGQVAAGAFADLRALRALHLDSNRLAEVRGDQLRGLGNLRHLILGNNQIRKVE 153

Qy          141 DEAFEDFLLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEGTFADLQKLA 200
                  ||: || |:||||||||||  |||::| :||||: |:|||||:||||||||  | ||
Db          154 SAAFDAFLSTVEDLDLSYNNLEALPWEAVGQMVNLNTLTLDHNLIDHIAEGTFVQLHKLV 213

Qy          201 RLDLTSNRLQKLPPDPIFARSQASALTATPFAP-PLSFSFGGNPLHCNCELLWLRRLERD 259
                |||:||||| ||||| :| |||       |  | ||: ||||||||||||||||||| |:
Db          214 RLDMTSNRLHKLPPDGLFLRSQGGG----PKPPTPLTVSFGGNPLHCNCELLWLRRLTRE 269

Qy          260 DDLETCGSPGGLKGRYFWHVREEEFVCEPPLIT-QHTHKLLVLEGQAATLKCKAIGDPSP 318
                |||||| :|  |  |||| : ||||:||||||| |   : ||:|||| :|:|:|:||| |
Db          270 DDLETCATPEHLTDRYFWSIPEEEFLCEPPLITRQAGGRALVVEGQAVSLRCRAVGDPEP 329

Qy          319 LIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMVEVSIVQ 378
                ::|||||| ||:|||||| |  :||||: ||| :||| |||||:|||||||| ||| :|
Db          330 VVHWVAPDGRLLGNSSRTRVRGDGTLDVTITTLRDSGTFTCIASNAAGEATAPVEVCVVP 389

Qy          379 LPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALV 438
                || :       ||| :    :|    :     |  |     |  || :: :|:|::| |:
Db          390 LPLM-------APPPAAPPPLTEPGSSDIATPGRPGAN--DSATERRLVAAELTSSSVLI 440

Qy          439 KWSVSKSAPRVKMYQLQYNCSDDEVLIY 466
                :|   :  | ::|||:||| | |: |:|
Db          441 RWPAQRPVPGIRMYQVQYNSSADDSLVY 468
```

```
RESULT 8
Q8TBS9
ID   Q8TBS9      PRELIMINARY;      PRT;   542 AA.
AC   Q8TBS9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to KIAA1484 protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC025310; AAH25310.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   542 AA;  57256 MW;  F5B1D5A46F326A51 CRC64;

  Query Match             26.0%;  Score 1063.5;  DB 4;  Length 542;
  Best Local Similarity   41.2%;  Pred. No. 3e-74;
  Matches  252;  Conservative   86;  Mismatches  165;  Indels  109;  Gaps   18;

Qy          217 IFARSQASALTATPFAP-PLSFSFGGNPLHCNCELLWLRRLERDDDLETCGSPGGLKGRY 275
                :| ||| :     |  | ||: ||||||||||||||||||| |:|||||| :|  |  ||
Db            1 LFLRSQGTG----PKPPTPLTVSFGGNPLHCNCELLWLRRLTREDDLETCATPEHLTDRY 56

Qy          276 FWHVREEEFVCEPPLIT-QHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSS 334
                || : ||||:||||||| |   : ||:|||| :|:|:|:||| |::|||||| ||:||||
Db           57 FWSIPEEEFLCEPPLITRQAGGRALVVEGQAVSLRCRAVGDPEPVVHWVAPDGRLLGNSS 116

Qy          335 RTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKS 394
                || |  :||||: ||| :||| |||||:|||||||| ||| :| || :       ||| :
Db          117 RTRVRGDGTLDVTITTLRDSGTFTCIASNAAGEATAPVEVCVVPLPLM-------APPPA 169

Qy          395 RLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQL 454
                    :|    :     |  |     |  || :: :|:|: | |::|   :  | ::|||:
Db          170 APPPLTEPGSSDIATPGRPGAN--DSAAERRLVAAELTSNSVLIRWPAQRPVPGIRMYQV 227

Qy          455 QYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFF 514
                ||| | |: |:|||||:::: |:||:| :|  |||||||::|| || | || :||| ||
Db          228 QYNSSVDDSLVYRMIPSTSQTFLVNDLAAGRAYDLCVLAVYDDGATALPATRVVGCVQFT 287

Qy          515 TKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRYKV-----CNHEAPSKMA 569
                |  |   |: : :  ||||||: |||:|||::|||||:||:||||           |:
Db          288 TAGDPAPCRPLRAHFLGGTMIIAIGGVIVASVLVFIVLLMIRYKVYGDGDSRRVKGSRSL 347

Qy          570 AAVSNVYSQTNGAQPPPPSSAPAGAPP-------------QGPPKVVVRNELLDFTASLA 616
                  ||:| ||||||      :| | | |             |  | | |           |
Db          348 PRVSHVCSQTNGA---GTGAAQAPALPAQDHYEALREVESQAAPAVAVE----------A 394

Qy          617 RASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAFA-SLDLKSQRKEELLD 675
                :| :: ::|:                      |:  | | :|  | || |    :    :
Db          395 KAMEAETASA---------------------EPEVVLGRSLGGSATSLCLLPSEETSGEE 433

Qy          676 SRTPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLP----LEGKAKRSHSFDMGDFAA 731
                ||         :| |    |   | || :   :|  :|       : :: :||| ||:
Db          434 SR---------AAVGPRRSRSGALEPPTSAPPTLALVPGGAAARPRPQQRYSFD-GDY-- 481

Qy          732 AAAGGVVPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARG--------------T 777
                    | :   :| ||:     |||  |      |  :    || |
Db          482 ----GALFQSHSYPRRARR--TKRHRST-----PHLDGAGGGAAGEDGDLGLGSARACLA 530

Qy          778 FGSSEWVMESTV 789
                | |:||::||||
Db          531 FTSTEWMLESTV 542

RESULT 9
```

```
Q96C50
ID   Q96C50      PRELIMINARY;      PRT;   450 AA.
AC   Q96C50;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical 47.4 kDa protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014678; AAH14678.1; -.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00041; fn3; 1.
DR   Pfam; PF00047; ig; 1.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   450 AA;  47417 MW;  93CEE7B44D2C1FB6 CRC64;

  Query Match             18.9%;  Score 775;  DB 4;  Length 450;
  Best Local Similarity   38.8%;  Pred. No. 6.7e-52;
  Matches  188;  Conservative   75;  Mismatches  138;  Indels   84;  Gaps   13;

Qy        297 KLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGA 356
              : ||:|||| :|:|:|:||| |::|||||| ||:|||||| |  :||||: ||| :|||
Db          6 RALVVEGQAVSLRCRAVGDPEPVVHWVAPDGRLLGNSSRTRVRGDGTLDVTITTLRDSGT 65

Qy        357 FTCIAANAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGE 416
              |||||:|||||||| ||| :| || :       ||| :    :|    :     |  |
Db         66 FTCIASNAAGEATAPVEVCVVPLPLM-------APPPAAPPPLTEPGSSDIATPGRPGAN 118

Qy        417 PPKSPPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAF 476
                 |  || :: :|:|: | |::|   :  | ::|||:||| | ||: |:|||||:::: |
Db        119 --DSAAERRLVAAELTSNSVLIRWPAQRPVPGIRMYQVQYNSSVDDSLVYRMIPSTSQTF 176

Qy        477 VVNNLVSGTGYDLCVLAMWDDTATTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMIL 536
              :||:| :|  |||||||::|| || | || :||| || |  |   |: : :  ||||||:
Db        177 LVNDLAAGRAYDLCVLAVYDDGATALPATRVVGCVQFTTAGDPAPCRPLRAHFLGGTMII 236

Qy        537 VIGGIIVATLLVFIVILMVRYKV-----CNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAP 591
               |||:|||::|||||:||:||||           |:    ||:| ||||||      :|
Db        237 AIGGVIVASVLVFIVLLMIRYKVYGDGDSRRVKGSRSLPRVSHVCSQTNGA---GTGAAQ 293

Qy        592 AGAPP-------------QGPPKVVVRNELLDFTASLARASDSSSSSSLGSGEAAGLGRA 638
              | | |             |  | | |           |:| :: ::|:
Db        294 APALPAQDHYEALREVESQAAPAVAVE----------AKAMEAETASA------------ 331

Qy        639 PWRIPPSAPRPKPSLDRLMGAFA-SLDLKSQRKEELLDSRTPAGRGAGTSARGHHSDREP 697
                        |:  | | :|  | || |    :    :||         :| |    |
Db        332 ---------EPEVVLGRSLGGSATSLCLLPSEETSGEESR---------AAVGPRRSRSG 373

Qy        698 LLGPPAARARSLLPLP----LEGKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNIWT 753
               | || :   :|  :|       : :: :||| ||:      | :   :| ||:     |
Db        374 ALEPPTSAPPTLALVPGGAAARPRPQQRYSFD-GDY------GALFQSHSYPRRARR--T 424

Qy        754 KRSLS 758
              ||  |
Db        425 KRHRS 429

RESULT 10
Q9BWJ0
ID   Q9BWJ0      PRELIMINARY;      PRT;   324 AA.
AC   Q9BWJ0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical 33.2 kDa protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EYE;
RA   Strausberg R.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COLON;
RA   Strausberg R.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC000207; AAH00207.1; -.
DR   EMBL; BC014040; AAH14040.1; -.
DR   EMBL; BC015581; AAH15581.1; -.
DR   InterPro; IPR003961; FN_III.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   324 AA;  33192 MW;  C84CCC9A811E877D CRC64;

  Query Match             16.0%;  Score 653;  DB 4;  Length 324;
  Best Local Similarity   34.1%;  Pred. No. 1.3e-42;
  Matches  161;  Conservative   54;  Mismatches  107;  Indels  150;  Gaps   10;

Qy        320 IHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMVEVSIVQL 379
              :||| |||||||||||   : ||||:| :| : |:| :|||| | |||||| ||: :: |
Db          1 MHWVGPDDRLVGNSSRARAFPNGTLEIGVTGAGDAGGYTCIATNPAGEATARVELRVLAL 60

Qy        380 PHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVK 439
              ||  ||::    |    |||  |::|:  | |:   ||       || |:||| || ||
Db         61 PHGGNSSAEGGRPGP--SDIAASARTAAEGEGTLESEP-------AVQVTEVTATSGLVS 111

Qy        440 WSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTA 499
              |   : |  | |:|:||| |:|| ||||::|||:  |:: :|| |  ||||:||:
Db        112 WGPGRPADPVWMFQIQYNSSEDETLIYRIVPASSHHFLLKHLVPGADYDLCLLALSPAAG 171

Qy        500 -TTLTATNIVGCAQFFTKADYPQCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRYK 558
               : |||| ::||| | |    | | :: : :||||: : :||::|| |||| | |:|| :
Db        172 PSDLTATRLLGCAHFSTLPASPLCHALQAHVLGGTLTVAVGGVLVAALLVFTVALLVRGR 231

Qy        559 -VCNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQGPPKVVVRNELLDFTASLAR 617
                 |   | |:    |:| |||||     ||  |   ||: ||
Db        232 GAGNGRLPLKL----SHVQSQTNGG----PSPTPKAHPPRSPP----------------- 266

Qy        618 ASDSSSSSSLGSGEAAGLGRAPWRIPPSAPRPKPSLDRLMGAFASLDLKSQRKEELLDSR 677
                                           |||
Db        267 -----------------------------PRP---------------------------- 269

Qy        678 TPAGRGAGTSARGHHSDREPLLGPPAARARSLLPLPLEGKAKRSHSFDMGDFAAAAAGGV 737
                                                       :|| | |:||
Db        270 -----------------------------------------QRSCSLDLGD--------- 279

Qy        738 VPGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSEWVMESTV 789
                | |   |::   | :|| ||:| ||        | ||  ||:| : || |
Db        280 -AGCYGYARRLGGAWARRSHSVHGGLLG------AGCRGVGGSAERLEESVV 324

RESULT 11
Q9HCJ2
ID   Q9HCJ2      PRELIMINARY;      PRT;   640 AA.
AC   Q9HCJ2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   KIAA1580 protein (Fragment).
GN   KIAA1580.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20450683; PubMed=10997877;
RA   Nagase T., Kikuno R., Nakayama M., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes.
RT   XVIII. The complete sequences of 100 new cDNA clones from brain which
RT   code for large proteins in vitro.";
RL   DNA Res. 7:273-281(2000).
DR   EMBL; AB046800; BAB13406.1; -.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 9.
DR   Pfam; PF01463; LRRCT; 1.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00370; LRR; 6.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 8.
KW   Immunoglobulin domain.
FT   NON_TER       1      1
SQ   SEQUENCE   640 AA;  71949 MW;  6058974872636838 CRC64;

  Query Match             10.6%;  Score 435.5;  DB 4;  Length 640;
  Best Local Similarity   24.5%;  Pred. No. 3.1e-25;
  Matches  163;  Conservative   87;  Mismatches  238;  Indels  177;  Gaps   19;

Qy           2 ETLLGGLLAFGMAFAV----VDACPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVEL 57
               : ||  |||  :           ||  | | |    :  :|  | |  ||  |   |  |
Db          24 DPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKV--ICVRKNLREVPDGISTNTRLL 81

Qy          58 RLGGNFIIHISRQDFANMTGLVDLTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGE 117
                |  | |  |    | ::  |  | |||| |  |:  :|  | :| :| |  ||| ::
Db          82 NLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLANLNTLELFDNRLTTIPN 141

Qy         118 DTLRGLVNLQHLIVNNN-------------------------QLGGIADEAFEDF----- 147
                    |  |: | : ||                         :|  |:: |||
Db         142 GAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLSYISEGAFEGLSNLRY 201

Qy         148 ----------------LLTLEDLDLSYNNLHGLPWDSVRRMVNLHQLSLDHNLLDHIAEG 191
                               |: |::|||| |:|  :   | : :::| :| :  : :  |
Db         202 LNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKLWMIQSQIQVIERN 261

Qy         192 TFADLQKLARLDLTSNRLQKLPPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELL 251
                | :|| |  ::|  | |  || | :|          ||           || :|||::|
Db         262 AFDNLQSLVEINLAHNNLTLLPHD-LF----------TPLHHLERIHLHHNPWNCNCDIL 310

Qy         252 WLRRLERD------DDLETCGSPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQA 305
               ||    :|           | :|  |||||   : :  | |  |:| :    | | || |
Db         311 WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMA 370

Qy         306 ATLKCKAIGDPSPLIHWVAPDDRLVGNSS---RTAVYDNGTLDIFITTSQDSGAFTCIAA 362
               | |||:|       : |: |:  :: : :   | ||  :|||:    | ||:| :||: :
Db         371 AELKCRA-STSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVS 429

Qy         363 NAAGEATAMVEVSIVQLPHLSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPP 422
               |: |  ||   :         | |: |  | |  | :|
Db         430 NSVGNTTASATL---------NVTAATTTPFSYFSTVT---------------------- 458

Qy         423 ERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLV 482
                                                         :  | |: ::|   :| |
Db         459 ------------------------------------------VETMEPSQDEARTTDNNV 476

Qy         483 SGTGYDLCVLAMWDDT--ATTLTATNIVGCAQFFT------KADYPQCQSMHSQILGGTM 534
                 |      :  |: |   |:||  :     : ||       :  |       :::  |
Db         477 GPT-----PVVDWETTNVTTSLTPQSTRSTEKTFTIPVTDINSGIPGI----DEVMKTTK 527

Qy         535 ILVIGGIIVATLLVFIVILMVRYKV------CNHEAPSKMAAAVSNVYSQTNGAQP---- 584
               |::  |  ||  |:  |:|:: ||:       || ||::    : ||  :  |  |
Db         528 III--GCFVAITLMAAVMLVIFYKMRKQHHRQNHHAPTR-TVEIINVDDEITGDTPMESH 584

Qy         585 -PPPS 588
                | |:
Db         585 LPMPA 589


RESULT 12
Q9BZ20
ID   Q9BZ20      PRELIMINARY;      PRT;   606 AA.
AC   Q9BZ20;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   BA438B23.1 (Neuronal leucine-rich repeat protein) (CDNA FLJ31810 fis,
DE   clone NT2RI2009289, weakly similar to carboxypeptidase N 83 kDa
DE   chain).
GN   BA438B23.1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Babbage A.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA   Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA   Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA   Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA   Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA   Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA   Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA   Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL353746; CAC22713.1; -.
DR   EMBL; AK056372; BAB71167.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 10.
DR   Pfam; PF01463; LRRCT; 1.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 10.
```

```
KW   Immunoglobulin domain.
SQ   SEQUENCE   606 AA;  68065 MW;  CB608E281B066B9D CRC64;

  Query Match             10.2%;  Score 418.5;  DB 4;  Length 606;
  Best Local Similarity   26.0%;  Pred. No. 6.1e-24;
  Matches  125;  Conservative   67;  Mismatches  157;  Indels  131;  Gaps   12;

Qy         21 CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
              ||  | |   ::|:   |  : |: :|  |   |  | |  | :  :: ::| :   | :
Db         28 CPARCECSAQNKSVS--CHRRRLIAIPEGIPIETKILDLSKNRLKSVNPEEFISYPLLEE 85

Qy         81 LTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIA 140
              : || | |::::| :| :| :|||| |  |||  :      || ||  | :: |::  :
Db         86 IDLSDNIIANVEPGAFNNLFNLRSLRLKGNRLKLVPLGVFTGLSNLTKLDISENKIVILL 145

Qy        141 DEAFEDF-----------------------LLTLEDLDLSYNNLHGLPWDS---VRRMVN 174
              |  |:|                        ||:|| | |   ||  :| ::   :| :::
Db        146 DYMFQDLHNLKSLEVGDNDLVYISHRAFSGLLSLEQLTLEKCNLTAVPTEALSHLRSLIS 205

Qy        175 LH---------------------------------------------------------- 176
              ||
Db        206 LHLKHLNINNMPVYAFKRLFHLKHLEIDYWPLLDMMPANSLYGLNLTSLSVTNTNLSTVP 265

Qy        177 -----------QLSLDHNLLDHIAEGTFADLQKLARLDLTSNRLQKLPPDPI-------- 217
                          |:| :| :  |  | |:|| :|  | :   :|: : |
Db        266 FLAFKHLVYLTHLNLSYNPISTIEAGMFSDLIRLQELHIVGAQLRTIEPHSFQGLRFLRV 325

Qy        218 --FARSQASALTATPFAPPLS---FSFGGNPLHCNCELLWLRRLERDDDLE------TCG 266
                 :::    |    |: | :    |   ||| |:| |||:  |:|   |:       |
Db        326 LNVSQNLLETLEENVFSSPRALEVLSINNNPLACDCRLLWI--LQRQPTLQFGGQQPMCA 383

Qy        267 SPGGLKGRYFWHVREEE----FVCEPPLI----TQHTHKLLVLEGQAATLKCKAIGDPSP 318
               |  :: | |           | |: | |     ||   ||| |||   |:| | ||| |
Db        384 GPDTIRERSFKDFHSTALSFYFTCKKPKIREKKLQH---LLVDEGQTVQLECSADGDPQP 440

Qy        319 LIHWVAPDDRLVGNSS--RTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMVEVSI 376
              :| || |  | :   |  |  |  :|||:|     |||| : |||:||||  |    :::
Db        441 VISWVTPRRRFITTKSNGRATVLGDGTLEIRFAQDQDSGMYVCIASNAAGNDTFTASLTV 500

RESULT 13
Q96A85
ID   Q96A85      PRELIMINARY;      PRT;   649 AA.
AC   Q96A85;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical 72.3 kDa protein.
GN   BAG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang J.;
RL   Thesis (2000), Zhongshan Medical University, Guangzhou, China.
DR   EMBL; AJ297858; CAC82651.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 8.
DR   Pfam; PF01463; LRRCT; 1.
DR   Pfam; PF01462; LRRNT; 1.
DR   SMART; SM00013; LRRNT; 1.
KW   Hypothetical protein; Signal.
FT   SIGNAL        1     38       POTENTIAL.
SQ   SEQUENCE   649 AA;  72313 MW;  9C11C5ABC7E536CF CRC64;

  Query Match              9.4%;  Score 385;  DB 4;  Length 649;
  Best Local Similarity   23.8%;  Pred. No. 2.7e-21;
```

```
  Matches  152;  Conservative   87;  Mismatches  239;  Indels  162;  Gaps   20;

Qy         21 CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
              ||  | | |    :  :|  :||  ||  |   |  | |  | |  |    | ::  |
Db         46 CPSVCSCSNQFSKV--VCTRRGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEV 103

Qy         81 LTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIA 140
              | | ||:|  |:  :|  | || :| |  | |  :       |  |: | : || :  |
Db        104 LQLGRNSIRQIEVGAFNGLASLNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIP 163

Qy        141 DEAFE-----------------------DFLLTLEDLDLSYNNLHGLPWDSVRRMVNLH 176
                ||                         : |  |: |:|   |:  :|  ::  :| |
Db        164 SYAFNRVPSLMRLDLGELKKLEYISEGAFEGLFNLKYLNLGMCNIKDMP--NLTPLVGLE 221

Qy        177 QLSLDHNLLDHIAEGTFADLQKLAR--------------------LDLTSNRLQKLPPDP 216
              :| :  |    |  |:|  |  | :                    |:|  | |  || |
Db        222 ELEMSGNHFPEIRPGSFHGLSSLKKLWVMNSHERNAFDGLASLVELNLAHNNLSSLPHD- 280

Qy        217 IFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLERD---DDLETCG---SPGG 270
              :|          ||    :      || :|:|::|||    |:    :   ||   :|
Db        281 LF----------TPLRYLVELHLHHNPWNCDCDILWLAWWLREYIPTNSTCCGRCHAPMH 330

Qy        271 LKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLV 330
              ::|||   | :  | |  | |      | : ||: | |||:    |   : |: |:  ::
Db        331 MRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAELKCRT--PPMSSVKWLLPNGTVL 388

Qy        331 GNSS---RTAVYDNGTLDIFITTSQDSGAFTCIAANAAG--EATAMVEVSIVQLPHLSNS 385
               ::|   | :| ::|||:       |:| :||:  | ||   |:| : ||  :| : ||
Db        389 SHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSNASAYLNVSTAEL-NTSNY 447

Qy        386 TSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVSKS 445
              :  |                                     :  | |  |
Db        448 SFFT-------------------------------------TVTVETTEIS---------- 461

Qy        446 APRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLTAT 505
                            :|    |: :| ::           |||           | | : |
Db        462 -------------PEDTTRKYKPVPTTS-----------TGY---------QPAYTTSTT 488

Qy        506 NIVGCAQFFTKADYP------QCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRYKV 559
               ::   :   :   |      : |:   :::  | |::  |  ||  |:   :|:| ||:
Db        489 VLIQTTRVPKQVAVPATDTTDKMQTSLDEVMKTTKIII--GCFVAVTLLAAAMLIVFYKL 546

Qy        560 -CNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQG 598
                 |:  | : || :    | :    |  :|| | | | |
Db        547 RKRHQQRSTVTAARTVEIIQVD-EDIPAATSAAATAAPSG 585

RESULT 14
Q8VI35
ID   Q8VI35      PRELIMINARY;      PRT;   648 AA.
AC   Q8VI35;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Brain tumor-associated protein MBAG1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/J;
RA   Wang J., Bin L., Li G.;
RT   "Brain-specific gene, downregulated in brain tumor.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF300458; AAL67671.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
```

```
DR   Pfam; PF00560; LRR; 8.
DR   Pfam; PF01463; LRRCT; 1.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00370; LRR; 5.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 6.
SQ   SEQUENCE   648 AA;  72215 MW;  E7E2EE50FC8F4DDD CRC64;

  Query Match             9.4%;  Score 383.5;  DB 11;  Length 648;
  Best Local Similarity   23.4%;  Pred. No. 3.6e-21;
  Matches  151;  Conservative   88;  Mismatches  238;  Indels  169;  Gaps   20;

Qy           21 CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
                ||  | | |    :  :|  :||  ||  |   |  | |  | |  |    | ::  |
Db           45 CPSVCSCSNQFSKV--VCTRRGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEV 102

Qy           81 LTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIA 140
                | | ||:|  |:  :|  | || :| |  | |  :       |  |: | : || :  |
Db          103 LQLGRNSIRQIEVGAFNGLASLNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIP 162

Qy          141 DEAFE------------------------DFLLTLEDLDLSYNNLHGLPWDSVRRMVNLH 176
                  ||                         : |  |: |:|   |:  :|  ::  :| |
Db          163 SYAFNRVPSLMRLDLGELKKLEYISEGAFEGLFNLKYLNLGMCNIKDMP--NLTPLVGLE 220

Qy          177 QLSLDHNLLDHIAEGTFADLQKLAR--------------------LDLTSNRLQKLPPDP 216
                :| :  |    |  |:|  |  | :                    |:|  | |  || |
Db          221 ELEMSGNHFPEIRPGSFHGLSSLKKLWVMNSHERNAFDGLASLVELNLAHNNLSSLPHD- 279

Qy          217 IFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLERD---DDLETCG---SPGG 270
                :|          ||    :      || :|:|::|||    |:    :   ||   :|
Db          280 LF----------TPLRYLVELHLHHNPWNCDCDILWLAWWLREYIPTNSTCCGRCHAPMH 329

Qy          271 LKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPDDRLV 330
                ::|||   | :  | |  | |      | : | : | ||||:    |   : |: |:  ::
Db          330 MRGRYLVEVDQAAFQCSAPFIMDAPRDLNISEDRMAELKCRT--PPMSSVKWLLPNGTVL 387

Qy          331 GNSS---RTAVYDNGTLDIFITTSQDSGAFTCIAANAAG--EATAMVEVSIVQL--PHLS 383
                 ::|   | :| ::|||:       |:| :||:  | ||   |:| : ||  :|  |: |
Db          388 SHASRHPRISVLNDGTLNFSRVLLIDTGVYTCMVTNVAGNSNASAYLNVSSAELNTPNFS 447

Qy          384 NSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWSVS 443
                  |: |                                        | |  |
Db          448 FFTTVTV---------------------------------------ETTEIS-------- 460

Qy          444 KSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATTLT 503
                                :|    |: :| ::           |||           | | :
Db          461 ---------------PEDITRKYKPVPTTS-----------TGY---------QPAYTTS 485

Qy          504 ATNIVGCAQFFTKADYP------QCQSMHSQILGGTMILVIGGIIVATLLVFIVILMVRY 557
                 | ::   :   :   |      : |:   :::  | |::  |  ||  |:   :|:| |
Db          486 TTVLIQTTRVPKQVPVPSTDTTDKMQTSLDEVMKTTKIII--GCFVAVTLLAAAMLIVFY 543

Qy          558 KV-CNHEAPSKMAAA----VSNVYSQTNGAQPPPPSSAPAGAPPQG 598
                |:   |:  | : ||    :  |      | |   ::||:|   :|
Db          544 KLRKRHQQRSTVTAARTVEIIQVDEDIPAAAPAAATAAPSGVSGEG 589

RESULT 15
Q9HBW1
ID   Q9HBW1      PRELIMINARY;      PRT;   653 AA.
AC   Q9HBW1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Brain tumor associated protein NAG14.
GN   NAG14.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang J., Bin L., Jiang N., Li G.;
RT   "Homo sapiens brain-specific gene (BAG), downregulated in brain tumor,
RT   mRNA.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF196976; AAG28019.2; -.
DR   HSSP; P23945; 1XUN.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00560; LRR; 8.
DR   Pfam; PF01463; LRRCT; 1.
DR   Pfam; PF01462; LRRNT; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00370; LRR; 3.
DR   SMART; SM00082; LRRCT; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 7.
KW   Immunoglobulin domain.
SQ   SEQUENCE   653 AA;  72717 MW;  38159C81F6850E37 CRC64;

  Query Match             9.4%;  Score 383;  DB 4;  Length 653;
  Best Local Similarity   23.6%;  Pred. No. 4e-21;
  Matches  152;  Conservative   87;  Mismatches  239;  Indels  166;  Gaps   20;

Qy           21 CPKYCVCQNLSESLGTLCPSKGLLFVPPDIDRRTVELRLGGNFIIHISRQDFANMTGLVD 80
                ||  | | |    :  :|  :||  ||  |   |  | |  | |  |    | ::  |
Db           46 CPSVCSCSNQFSKV--VCTRRGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEV 103

Qy           81 LTLSRNTISHIQPFSFLDLESLRSLHLDSNRLPSLGEDTLRGLVNLQHLIVNNNQLGGIA 140
                | | ||:|  |:  :|  | || :| |  | |  :       |  |: | : || :  |
Db          104 LQLGRNSIRQIEVGAFNGLASLNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIP 163

Qy          141 DEAFE------------------------DFLLTLEDLDLSYNNLHGLPWDSVRRMVNLH 176
                  ||                         : |  |: |:|   |:  :|  ::  :| |
Db          164 SYAFNRVPSLMRLDLGELKKLEYISEGAFEGLFNLKYLNLGMCNIKDMP--NLTPLVGLE 221

Qy          177 QLSLDHNLLDHIAEGTFADLQKLAR------------------------LDLTSNRLQKL 212
                :| :  |    |  |:|  |  | :                        |:|  | |  |
Db          222 ELEMSGNHFPEIRPGSFHGLSSLKKLWVMNSQVSLIERNAFDGLASLVELNLAHNNLSSL 281

Qy          213 PPDPIFARSQASALTATPFAPPLSFSFGGNPLHCNCELLWLRRLERD---DDLETCG--- 266
                | | :|          ||    :      || :|:|::|||    |:    :   ||
Db          282 PHD-LF----------TPLRYLVELHLHHNPWNCDCDILWLAWWLREYIPTNSTCCGRCH 330

Qy          267 SPGGLKGRYFWHVREEEFVCEPPLITQHTHKLLVLEGQAATLKCKAIGDPSPLIHWVAPD 326
                :|  ::|||   | :  | |  | |      | : ||: | |||:    |   : |: |:
Db          331 APMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAELKCRT--PPMSSVKWLLPN 388

Qy          327 DRLVGNSS---RTAVYDNGTLDIFITTSQDSGAFTCIAANAAG--EATAMVEVSIVQLPH 381
                  :: ::|   | :| ::|||:       |:| :||:  | ||   |:| : ||  :| :
Db          389 GTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSNASAYLNVSTAEL-N 447

Qy          382 LSNSTSRTAPPKSRLSDITGSSKTSRGGGGSGGGEPPKSPPERAVLVSEVTTTSALVKWS 441
                 || :  |                                     :  | |  |
Db          448 TSNYSFFT------------------------------------TVTVETTEIS------ 465

Qy          442 VSKSAPRVKMYQLQYNCSDDEVLIYRMIPASNKAFVVNNLVSGTGYDLCVLAMWDDTATT 501
                                  :|    |: :| ::           |||           | |
Db          466 -----------------PEDTTRKYKPVPTTS-----------TGY---------QPAYT 488

Qy          502 LTATNIVGCAQFFTKADYP------QCQSMHSQILGGTMILVIGGIIVATLLVFIVILMV 555
```

```
                 : | ::   :  :   |      : |:   :::  | |::  |  || |:   :|:|
Db          489 TSTTVLIQTTRVPKQVAVPATDTTDKMQTSLDEVMKTTKIII--GCFVAVTLLAAAMLIV 546

Qy          556 RYKV-CNHEAPSKMAAAVSNVYSQTNGAQPPPPSSAPAGAPPQG 598
                 ||:   |:  | : || :   | :    |  :|| | | | |
Db          547 FYKLRKRHQQRSTVTAARTVEIIQVD-EDIPAATSAAATAAPSG 589

Search completed: June  8, 2003, 21:30:25
Job time : 69 secs
```